

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 29, 2005, 18:47:48 ; Search time 1178 Seconds
(without alignments)
19529.306 Million cell updates/sec

Title: US-10-613-728-1

Perfect score: 5735
1 ggcattctgcaaggtcacaca.....ccataagagttcgtcgcac 5735

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gsa1:*
9: gb_gsa2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	534.4	9.3	613	8	AZ269661 RPCI-23-4
C 2	499.4	8.7	529	8	BH073204 RPCI-24-3
C 3	473.8	8.3	496	8	AZ300332 RPCI-23-1
C 4	305.6	5.3	312	8	AZ457684 IM0261E12
5	305.2	5.3	469	8	BH285252 CH230-118
6	266.4	4.6	447	8	AZ729312 RPCI-24-7
7	236	4.1	447	5	BY000251 BY000251
8	225.2	3.9	616	2	BF525096 UI-R-AC0-
9	145.4	2.5	370	8	AZ407663 IM0178C01
10	133.4	2.3	390	5	BY000731 BY000731
11	133.4	2.3	409	5	BY061731 BY061731
12	133.4	2.3	415	5	BY083328 BY083328
13	133.4	2.3	739	6	CB596657 AGENCOURT
14	133.4	2.3	741	6	CB599099 AGENCOURT
15	133.4	2.3	783	6	CA490238 AGENCOURT
16	133.4	2.3	843	6	CA460851 AGENCOURT
17	133.4	2.3	843	6	CA495692 AGENCOURT
18	133.4	2.3	850	6	CA490440 AGENCOURT
19	133.4	2.3	917	6	CA490678 AGENCOURT
20	133.4	2.3	920	6	CA490677 AGENCOURT
21	132.8	2.3	871	6	CA490555 AGENCOURT
22	132.4	2.3	363	5	BY083663 BY083663
23	131.8	2.3	881	6	CB951696 AGENCOURT
24	130	2.3	253	5	BK632021 BK632021

C 25	130	2.3	459	1	AI642327	AI642327 vna6f10.Y
26	129.6	2.3	378	5	BY064144	BY064144 BY064144
27	129.2	2.3	224	5	BY000659	BY000659 BY000659
C 28	129	2.2	707	1	AA855323	AA855323 w71a12.Y
29	127	2.2	349	1	AI892571	AI892571 mt79f09.Y
30	126.8	2.2	412	5	BY435020	BY435020 BY435020
31	126.8	2.2	468	5	BY414811	BY414811 BY414811
32	126.2	2.2	368	5	BY000397	BY000397 BY000397
33	125.2	2.2	368	2	BF472804	BF472804 UI-M-B20-
34	125.2	2.2	385	5	BY422669	BY422669 BY422669
C 35	125.2	2.2	428	5	BF462620	BF462620 UI-M-CGDP
36	124.8	2.2	399	5	BY432934	BY432934 BY432934
37	123.8	2.2	405	5	BY419314	BY419314 BY419314
38	123.8	2.2	425	5	BY427154	BY427154 BY427154
39	123.6	2.2	405	5	BY436012	BY436012 BY436012
C 40	123	2.1	426	1	AA140343	AA140343 mq89f06.Y
C 41	123	2.1	473	1	AA919406	AA919406 v213604.Y
42	122.8	2.1	203	5	BY426544	BY426544 BY426544
43	122.8	2.1	246	1	AV244652	AV244652 AV244652
C 44	122.8	2.1	278	4	BM124558	BM124558 L0541G01-
45	122.8	2.1	313	5	BY415619	BY415619 BY415619

ALIGNMENTS

RESULT 1
LOCUS AZ269661/c 613 bp DNA linear GSS 26-JUL-2000
DEFINITION RPCI-23-444F20.TU RPCI-23 Mus musculus genomic clone
ACCESSION AZ269661
VERSION AZ269661.1 GI:9483276
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Bukaryocota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 613)
Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatman,S.,
Aktiret,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de
Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-23
Unpublished (1999)
Other_GSSs: RPCI-23-444F20.TV
COMMENT
TITLE JOURNAL
COMMENT
Contact: Shanying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Piter de Jong
(piterdejong.med.buflalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buflalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/cdb/bac/ends/mouse/bac_end_intro.html
plate: 444 row: F column: 20
Seq primer: SP6
Class: BAC ends.

FEATURES

source
1..613
Location/Qualifiers
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-23-444F20"
/sex="Female"
/lab_host="DH10B"
/clone_11b="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site 1:
EcorI, Site 2: EcorI, Female C57BL/6J mouse kidney and/or

	DEFINITION
RPCI-23-117N22.TV RPCI-23 Mus musculus genomic clone RPCI-23-117N22, genomic survey sequence.	
AZS00332	
GSS,	GI:9542117
Mus musculus (house mouse)	
Mus musculus	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 496) Zhao,S., Niemann,W., Feldbljum,T., Malek,J., Shatman,S., Akinet,B., Levins,M., McGinn,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P. and Frazer,C.M. Mus BAC End Sequences from Library RPCI-23 Unpublished (1998) Other_GSSs: RPCI-23-117N22_TV Contact: Shaying Zhao Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel.: 301 838 0200 Fax: 301 838 0208 Email: szhao@igf.org	
Claones are derived from the mouse BAC library RPCI-23. Por BAC library availability, please contact Pieter de Jong (pieterdejong.med.buffalo.edu) . Claones may be purchased from BACPac Resources (http://bacpac.med.buffalo.edu/orderingframe.htm) or from Resea ch Genetics (info@resgen.com). BAC end page: http://www.tigr.org/tdb/bac ends/mouse/bac_end_intro.html Plate: 117 row: N column: 22 Seq primer: SP6 Class: BAC ends.	

```

FEATURES
    source
        Location/Qualifiers
            1..496
                /organism="Mus musculus"
                /mol_type="genomic DNA"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="RPC1-23-117N22"
                /sex="Female"
                /lab_host="DH10B"
                /clone_lib="RPC1-23"
                /note="Organ: Kidney/Brain; Vector: pBAC3.6; Site_1:
                EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
                brain genomic DNA was isolated and partially digested
                with a combination of EcoRI and EcoRI Methyase. Site
                selected DNA was cloned into the pBAC3.6 vector at the
                EcoRI sites. The ligation products were transformed into
                DH10B electrocompetent cells (BRL Life Technologies)."

```

	ORIGIN
Query Match	8.3% Score 473.8; DB 8;
Best Local Similarity	98.4%; Pred. No. 7,1e-114;
Matches 469; Conservative	0; Mismatches 7; Indels 1; Gaps 1

Qy	3 359	CCCACTTAACCTTTCTCCCTCAGACCCAGAGATTAAACACCTCTGGCTTCCCTCTCC	3 312
Db	4 96	CCCACTTAATACCTTTCTCCCTCAGACCCAGAGATTAAACACCTCTGGCTTCCCTCTCC	4 37
Qy	3 319	ACCTCCCATCAGAGATGAGGGTTGCAGAGGAGGGTAAAACTACATGTCACAAATC	3 378
Db	4 36	ACCTCCCATCAGAGATGAGGGTTGCAGAGGAGGGTAAAACTACATGTCACAAATC	3 77
Qy	3 379	ATGGGCAAGATTAATGTGATCAGTATGTGTGTGAGGCAAGAAAGAAATCTGAGGCTTA	3 433
Db	3 76	ATGGGCAAGATTAATGTGATCAGTATGTGTGTGAGGCAAGAAAGAAATCTGAGGC-TAA	3 18
Qy	3 439	CTGGGTTAAATGTGTTAAAGTCTGTGTGCATGTGTGTGTCTGACTGAAAACGGGCAATGGC	3 492
Db	3 17	CTGGGTTAAATGTGTTAAAGTCTGTGTGCATGTGTGTGTCTGACTGAAAACGGGCAATGGC	258
Qy	3 499	TGTGACAGCTGTTCACTTCTGTGTGTGCGTAGGTTAACAGACTGCAGAGTTTGTGTGTTAAATGGC	3 555

Db	257	TGTCACGCTGTTCAAGTCTGTGCGTAGGATTAACAGACTGAGAGGTTGTGTAAATTAGC	198
OY	3559	CCAAAGCAAAATGGTGTGATATCCCTTCCATAGGTTTAAAGAGTTGGATGATGCGCTGCATC	3618
Db	197	CCAAAGCAAAATGGTGTGATTAACACTTTCATAGGTTTAAAGAGTTGGATGATGCGCTGCATC	138
OY	3619	TCAAGACCAATGGAATAATAGATGACACTCTATATGTGTCTTACGTAAAGTAGCAAG	3678
Db	137	TCAGGACCAATGGAATAATAGATGACACTCTATATGTGTCTTACGTAAAGTAGCAAG	78
OY	3679	GTCTTTGGAGACACCTGTCTAAGAGATGTGGCAACAGACTTACAGACATATCTGTAC	3738
Db	77	GTCTTTGGAGACACCTGTCTAAGAGATGTGGCAACAGACTTACAGACATATCTGTAC	18
OY	3739	AGAGTAAAGAGAGAGAG 3755	
Db	17	AGAGTAAAGAGAGAGAG 1	

RESULT 4	
AZ457684/c	
LOCUS	312 bp DNA linear GSS 04-OCT-2000
DEFINITION	1M0261E12F Mouse 10kb plasmid UUGCM library Mus musculus genomic
ACCESSION	clone UUGCM0261E12 F, genomic survey sequence.
VERSION	AZ457684
KEYWORDS	AZ457684.1 GI:10615809
SOURCE	GSS.
ORGANISM	Mus musculus (house mouse)
	Mus musculus

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 312)	Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C., Ismail, H., Longacre, S., Mahmood, M., Meenen, E., Pedersen, T., Rellay, M., Rose, M., Rose, R., Stokes, R., Tinney, A., von Niederhausen, A. and Wright D. Weiss, R.	Mus whole genome scaffolding with paired end reads from 10kb plasmid inserts	Unpublished (2000)	Contact: Robert B. Weiss

University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT
84112, USA
Tel.: 801 585 5606
Fax: 801 585 7177
Email: dduan@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0261 row: E column: 12
Seq primer: CGTTGTAACAACGACGGCAGT
Class: plasmid ends
High quality sequence stop: 312.
Location/Qualifiers
1. .312

```

/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="TUGCIM0261E12"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10Kb plasmid UGCGM library"
/notes="Vector: PMD42nv; Purified genomic DNA from M.
musculus C57BL/6J (male); was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA
was hydrodynamically sheared by repeated passage through a
0.005 inch orifice at constant velocity. The sheared DNA
was blunt end-repaired with T4 DNA polymerase and T4
polynucleotide kinase. Adaptor oligonucleotides were
ligated to the blunt ends in high molar excess. The
adapted DNA was purified and size-selected for a 9.5 to
10.5 kb range using preparative agarose gel

```

electrophoresis. Vector DNA was prepared from a derivative of pMD2 (g1473114[gb/AP12972.1]), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptor complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 5.3%; Score 305.6; DB 8; Length 312;
Best Local Similarity 98.7%; Pred. No. 2.8e-69;
Matches 308; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 605 GTCTGGGGGTGTAGAACTGACAGTACTTTTCAGTGGGAAAGTATGACCCCTCAG 664
DB 312 GTCTGGGGGTGTAGAACTGACAGTACTTTTCAGTGGGAAAGTATGACCCCTCAG 253
QY 665 CAGATGTAGTATGTCCCTTATGATCCCATCCAGGAGGTCTCTAAGAGACATGGGAT 724
DB 252 CAGATGTAGTATGTCCCTTATGATCCCATCCAGGAGGTCTCTAAGAGACATGGGAT 193
QY 725 GAGGATGTAGTATGTCCCTTATGATCCCATCCAGGAGGTCTCTAAGAGACATGGGAT 764
DB 192 GAGGATGTAGTATGTCCCTTATGATCCCATCCAGGAGGTCTCTAAGAGACATGGGAT 133
QY 785 ACTTAGCAGAGAGACAGTAACTTATCTTCTCCCTCCATCCCTCCAGAGACA 844
DB 132 ACTTAGCAGAGAGACAGTAACTTATCTTCTCCCTCCATCCCTCCAGAGACA 73
QY 845 CACCCCTGTGTGACATTTCTTCTTCAAGTCCCTCTGTGACTTCCATTTGC 904
DB 72 CACCCCTGTGTGACATTTCTTCTTCAAGTCCCTCTGTGACTTCCATTTGC 13
QY 905 AAGGCTTTTGC 916
DB 12 AAGGCTTTTGC 1

RESULT 5 469 bp DNA linear GSS 30-NOV-2001
LOCUS BH285252
DEFINITION CH230-118C1.TVB CHORI-230 Segment 1 Rattus norvegicus genomic clone
ACCESSION BH285252
VERSION BH285252.1 GI:17197658
KEYWORDS GSS.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 469)
AUTHORS Zhao,S., Shetty,J., Shatsman,S., Teegaye,G., Geer,K., Shvartsbeyn,A., Gebregeorgis,E., Overton,L., Russell,D., Chen,D., R1998,F., de Jong,P. and Fraser,C.M.
Rat BAC End Sequences from Library CHORI-230 EcoRI segment Unpublished (1999)
OTHER GSSs: CH230-118C1.TVB
CONTACT: Shaying Zhao
DEPARTMENT OF Eukaryotic Genomics
THE INSTITUTE FOR GENOMIC RESEARCH
9712 Medical Center Dr., Rockville, MD 20850, USA
TEL: 301 838 0200
FAX: 301 838 0208
EMAIL: szhao@igr.org

Clones are derived from the rat BAC library CHORI-230 (http://www.chori.org/bacpac/rat230.htm). For BAC library availability, please contact Pieter de Jong (pdejong@mail.cho.org). Clones may be purchased from BACPAC Resources (http://www.chori.org/bacpac/orering information.htm). BAC end page: http://www.tigr.org/cdb/bac_ends/rat/bac_end_intro.html
Plate: 118 row: C column: 1

Seq primer: SP6
Class: BAC ends.
Location/Qualifiers
FEATURES
source 1..469
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="BN/SENHsd/MCW"
/db_xref="taxon:10116"
/clone="CH230-118C1"
/sex="Female"
/cell_type="Brain"
/clone_11b="CHORI-230 Segment 1"
/note="Vector: pTARBAC2.1; Site_1: EcoRI; Site_2: EcoRI; CHORI-230 Rat (BN/SENHsd/MCW) BAC library produced by Pieter de Jong"

ORIGIN

Query Match 5.3%; Score 305.2; DB 8; Length 469;
Best Local Similarity 89.2%; Pred. No. 4.2e-69;
Matches 387; Conservative 0; Mismatches 38; Indels 9; Gaps 5;

QY 3787 GGAACCTGAGTGTGACCTGCAAGTGTGCTCTCCCTAGACATCATGACTTGTCTC 3846
DB 2 GGAACCTGAGTGTGACCTGCAAGTGTGCTCTCCCTAGACATCATGACTTGTCTC 61
QY 3847 TGGGAGCCAGCAGCTGTGAACTTCAAGTCTGAGAGATGAGAGGCTCCCTCAGCCTGA 3906
DB 62 TGGGAGCCAGCAGCTGTGAACTTCAAGTCTGAGAGATGAGAGGCTCCCTCAGCCTGA 119
QY 3907 AGCTATGCAATATGCCAGGTTGAAA--GGGGAAGAGAGAGCTGGATGGAGCTTGT 3964
DB 120 AGCTATGCAATATGCCAGGTTGAAA--GGGGAAGAGAGAGCTGGATGGAGCTTGT 179
QY 3965 GTGTGAGAGAGAGAGAGATTTAAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 4023
DB 180 GTGTGAGAGAGAGAGAGATTTAAGCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAG 229
QY 4024 TTCACTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 4080
DB 240 TTCACTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 299
QY 4081 GTGTGAG 4140
DB 300 GTGTGAG 359
QY 4141 GCCCAAG 4200
DB 360 GCCCAAG 418
QY 4201 GCAACCTTGGGAGC 4214
DB 419 GCAACCTTGGGAGC 432

RESULT 6 447 bp DNA linear GSS 25-JAN-2001
LOCUS AZ729312
DEFINITION RPCI-24-79L13.TVB RPCI-24 Mus musculus genomic clone RPCI-24-79L13,
ACCESSION AZ729312
VERSION AZ729312.1 GI:12486941
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 447)
AUTHORS Zhao,S., Nieman,W., Malek,J., Shatsman,S., Akintec,B., Levine,M., Teegaye,G., Geer,K., Krol,M., Shvartsbeyn,A., Gebregeorgis,E., Russell,D., de Jong,P. and Fraser,C.M.
Mouse BAC End Sequences from Library RPCI-24 Unpublished (1999)
CONTACT: Shaying Zhao

Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhoet@igrr.org

Clones are derived from the mouse BAC library RPCI-24. For BAC library availability, please contact Piere de Jong (pdejong@igrr.org). Clones may be purchased from BACPAC Resources (<http://www.chori.org/bacpac/orderingframe.html>). BAC end page: http://www.chori.org/tbld/bac_ends/mouse/bac_intro.html. Plate: 79 row: L column: 13
Seq primer: T7
Class: BAC ends.

FEATURES

Location/Qualifiers

1..447
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="RPCI-24-79L13"
/sex="Male"
/cell_type="Spleen/Brain"
/note="Vector: PTARBAC1; Site 1: BamHI; Site 2: BamHI; RPCI-24 Mouse BAC Library produced by Piere de Jong. The library was cloned in the PTARBAC1 cloning vector at the BamHI sites using MboI partially digested male C57BL/6J DNA."

ORIGIN

Query Match 4.6%; Score 266.4; DB 8; Length 447;
Best Local Similarity 97.6%; Pred. No. 8.3e-59; Indels 1; Gaps 1;
Matches 281; Conservative 0; Mismatches 6;

QY 1 GGATCCCTGACAGGTACACAAAGGCTCTCCACCAAGGAGGCTCTGATTCATTTCAATTCAGT 60
DB 161 GGATCCCTGACAGG-CATACATGATGCTCCACCAAGGAGGCTCTGATTCATTTCAATTCAGT 219
QY 61 TTCATGCTCTGTTCTTCACATGCTGCGCTCCCAAGAGCTAATTTGAGCTTTGTTTAT 120
DB 220 TTCATGCTCTGTTCTTCACATGCTGCGCTCCCAAGAGCTAATTTGAGCTTTGTTTAT 279
QY 121 TTCAAAGGCGCTGATGAGAGTATGCTGCTACCAAGCTCTAAGGTCGCCGCTA 180
DB 280 TTCAAAGGCGCTGATGAGAGTATGCTGCTGCTACCAAGGTCGCCGCTA 339
QY 181 AGCCCTCAGACCTGAGAGCTTTGCAACAGCCCTTTAGTGAAGAGCAATTAAGCAATTT 240
DB 340 AGCCCTCAGACCTGAGAGCTTTGCAACAGCCCTTTAGTGAAGAGCAATTAAGCAATTT 399
QY 241 TCCTTAAGCAAAATCTGCTCTGAGCTCTTCTCTCTGACCTCG 288
DB 400 TCCTTAAGCAAAATCTGCTCTGAGCTCTTCTCTCTGACCTCG 447

RESULT 7
BY000251/c 447 bp mRNA linear EST 06-DEC-2002

LOCUS
DEFINITION
CDNA clone 1020001103 5', mRNA sequence.

ACCESSION
BY000251
VERSION
BY000251.1 GI:26060500

KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 447)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nishikado, I., Osato, N., Saito, R., Suzuki, H., Yamana, K., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Balderelli, R., Hill, D.P., Bult, C.,

JOURNAL MEDLINE PUBMED

COMMENT

Hume, D.A., Quackenbush, J., Schriml, L.M., Kanapin, A., Matsuda, H., Batilov, S., Beisel, K.W., Blake, J.A., Bradt, D., Brucic, V., Chochoi, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavani, M.J., Pereira, G., Peocle, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramchandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C.A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyszewski, B., Yang, I., Yang, L., Yuan, Z., Zavalan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterson, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@gsr.riken.jp, URL: <http://genome.gsc.riken.jp/>

Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Konno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanishi, A., Watanabe, M. and Hayashizaki, Y. Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multichannel sequencer. Genome Res. 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse cDNAs.

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

FEATURES

source

Location/Qualifiers

1..447
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="1020001103"
/sex="male"
/tissue_type="heart"
/dev stage="adult"
/clone_lib="BMM132"
/note="Site 1: Sali; Site 2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia

REFERENCE
AUTHORS
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 370)

TITLE
JOURNAL
COMMENT
Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah
Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0178 row: C column: 01
Seq primer: CACACAGAAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 370.

FEATURES
source
1. 370
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UGGCM0178C01"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/clone_lib="Mouse 10kb plasmid UGGCM library"
/note="Vector: pMD229v. Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pMD42 (g14732114|g14732072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Query Match 2.5%; Score 145.4; DB 8; Length 370;
Best Local Similarity 96.1%; Pred. No. 1.1e-26;
Matches 149; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 4062 GGCTGGGTAGGAGAGGTGTGAGACGCTCTGTCTCTCTGATGCGCTGAGGCC 4121
Db 1 GGCTGGGTAGGAGAGGTGTGAGACGCTCTGTCTCTCTGATGCGCTGAGGCC 60
QY 4122 CTTTGGGAGAGAGATGTGCGCCAGAGACTTAAAAAAGGCGATGAGGCCAGAGGGCGAG 4181
Db 61 CTTTGGGAGAGAGATGTGCGCCAGAGACTTAAAAAAGGCGATGAGGCCAGAGGGCGAG 120
QY 4182 GGCACAGAGACTTTCATGGGCAACCTTGGGGCCC 4216
Db 121 GGCACAGAGACTTTCATGGGCAACCTTGGGGCCC 155

RESULT 10
BY000731

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
BY000731 390 bp mRNA linear EST 06-DEC-2002
CDNA clone 1020013103 5', mRNA sequence.
BY000731
BY000731.1 GI:26060980
EST.
Mus musculus (house mouse)

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 390)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaide, I., Oatso, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Schiml, L.M., Kanapin, A., Matsuda, H., Hume, D.A., Quackenbush, J., Schiml, L.M., Brad, D., Brucic, V., Chochia, C., Corbani, L.E., Cousins, S., Dalla, E., Dragani, T.A., Fletcher, C.F., Forrest, A., Frazer, K.S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gustincich, S., Hirokawa, N., Jackson, I.J., Jarvis, E.D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedziarski, R.M., King, B.L., Konagaya, A., Kurochkin, I.V., Lee, Y., Lenhard, B., Lyons, P.A., Maglott, D.R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W.J., Petrea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J.U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J.C., Reed, D.J., Reid, J., Ring, B.Z., Ringwald, M., Sanderlin, A., Schneider, C., Sempke, C.A., Setou, M., Shinada, K., Sultana, R., Takenaka, Y., Taylor, M.S., Teasdale, R.D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L.G., Wyszewski, B., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E.S., Rogers, J., Birney, E. and Hayashizaki, Y.
Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
22354683
12466851
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suenryo-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-research@riken.jp URL: http://genome.gsc.riken.jp/
Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirazaki, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kono, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Nunazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
source

```
Location/Qualifiers
1..390
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="1020013103"
/sex="male"
/tissue_type="heart"
/dev_stage="adult"
/lab_host="BNN132"
/clone_lib="RIKEN full-length enriched, adult male heart"
/note="Site_1: Sali; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5' GAGAGAGAGAGCGCGCCGCACTCGAGTTTCTTTTCTTTTAA 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. Second strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGAGAGATCCAGAGCTCAATTAATTAATTAACCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBlueScript KS(+) after bulk excision from Lambda FLC I."
```

ORIGIN

Query Match 2.3%; Score 133.4; DB 5; Length 390;

Best Local Similarity 95.8%; Pred. No. 1,7e-23; Matches 137; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 5289 TCCTGTGAGAGATCTAGATGGTCTCCAGCCTCTGACTCTCTGCTGCTGTC 5348

Db 10 TCCACCCAGAGATCTCGATGTGTCTCCAGCCTCTGCTGCTCTCTCTGCTGTC 69

Qy 5349 CTCCTCTGTCCAGCTGCGCACTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCA 5408

Db 70 CTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCA 129

Qy 5409 GGATTTCTGAAAAGTTAACG 5431

Db 130 GGATTTCTGAAAAGTTAACG 152

RESULT 11

LOCUS BY061731

DEFINITION BY061731 RIKEN full-length enriched, pooled tissues, 16 days

ACCESSION BY061731

VERSION BY061731.1 GI:26181469

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

AUTHORS 1 (bases 1 to 409)

Okazaki, Y., Furuno, M., Kaekawa, T., Adachi, J., Bono, H., Kondo, S., Nakai, T., Otsu, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bull, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Brady, D., Brusic, V., Chotha, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T., Giacchino, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S., Gusticich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawai, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Kongaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nageshima, T.,

TITLE

JOURNAL MEDLINE PUBLISHED

COMMENT

Numata, K., Okido, T., Pavan, M. J., Perrea, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takemura, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilting, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Komno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashimoto, M., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suenho-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gs.riken.jp, URL: <http://genome.gsc.riken.jp/> Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hirozane, T., Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Komno, H., Miyazaki, A., Murata, M., Nakamura, M., Nomura, K., Numazaki, R., Ohno, M., Sakai, K., Sakazume, N., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M., Waki, K., Watanuki, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001) Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000) RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000) Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001) cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES
source

```
Location/Qualifiers
1..409
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone_lib="RIKEN full-length enriched, pooled tissues, 16 days embryo, etc."
/note="pooled tissues (dev stage=16 days embryo, tissue_type=heart, sex=mix), (dev stage=17 days embryo, tissue_type=kidney, sex=mix), (dev stage=17 days embryo, tissue_type=stomach, sex=mix), (dev stage=17 days embryo, tissue_type=kidney, sex=mix), (dev stage=17 days pregnant, adult, tissue_type=monon, sex=female), (dev stage=13 days embryo, tissue_type=liver, sex=mix)"
```

ORIGIN

Query Match 2.3%; Score 133.4; DB 5; Length 409; Best Local Similarity 95.8%; Pred. No. 1,7e-23; Matches 137; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Oy	5289	TCGCGGTAGAGATCTCTGATTGGTCTCCAGACCTCAGCTACTCTCTCTCCGCCTGTGTC	5348
Db	15	TCCAACCAGAGATCTCTGATTGGTCTCCAGACCTCTCTACTCTCTCTCTCCGCTGTTC	74
Oy	5349	CTCTCTGTGCCAGCTGGCCCACTGTGGTGCTCTGTTCCAGCTGTGTGTCCACATCTTCTCA	5408
Db	75	CTCTCTCCGTCCAGCTGGCCCACTGTGGTGCTCTGTTCCAGCTGTGTGTCCACATCTTCTCA	134
Oy	5409	GGATTCTCTGAAAAAAGTTAACCG	5431
Db	135	GGATTCTCTGAAAAAAGTTAACCG	157
RESULT_12			
LOCUS	BY083328		
DEFINITION	BY083328 RIKEN full-length enriched, pooled tissues, adult spleen,		
VERSION	BY083328		
SOURCE	BY083328.1 GI:26193492		
ORGANISM	Mus musculus (house mouse)		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
	1 (bases 1 to 415)		
	Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,		
	Nitafido,I., Osato,N., Saito,R., Suzuki,H., Yamamaka,I.,		
	Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,		
	Schombach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bull,C.,		
	Hume,P.A., Quackenbush,J., Schirral,L.M., Kanapin,A., Matsuda,H.,		
	Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,		
	Chochka,C., Corbani,L.E., Cousins,S., Daila,E., Dragalin,T.A.,		
	Fletcher,C.F., Forrest,A., Frazier,K.S., Gaasterland,T.,		
	Gairiboldi,M., Glasi,C., Godzik,A., Gough,J., Grimmond,S.,		
	Guindlich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,		
	Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,		
	Kurochkin,I.V., Lee,Y., Lehhard,B., Lyons,P.A., Meglott,D.R.,		
	Mutatsis,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,		
	Nunata,K., Okido,T., Pavan,W.J., Perlea,G., Pesole,G.,		
	Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramchandran,S.,		
	Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,		
	Sandelin,A., Schneider,C., Semple,C.A., Secou,M., Shmada,K.,		
	Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,		
	Veraardo,R., Wagner,L., Wahlesedec,C., Wang,Y., Watanabe,Y.,		
	Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,		
	Yang,L., Yuan,Z., Zavoian,M., Zhu,Y., Zimmer,A., Carninci,P.,		
	Hayatsu,N., Hirozane-Kishikawa,T., Komno,H., Nakamura,M.,		
	Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,		
	Aizawa,K.T., Fukuda,S., Hara,A., Hashizume,W., Imocani,K., Ihnii,Y.,		
	Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Saesaki,D., Shibata,K.,		
	Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,		
	Rogers,J., Birney,E. and Hayashizaki,Y.		
TITLE	Analysis of the mouse transcriptome based on functional annotation		
JOURNAL MEDLINE PUBMED COMMENT	of 60,770 full-length cDNAs Nature 420, 563-573 (2002) Contact: Yoshihide Hayashizaki Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute The Institute of Physical and Chemical Research (RIKEN) 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan Tel.: 81-45-503-9222 Fax: 81-45-503-9216 Email: genome-res@gsc.riken.jp, URL:http://genome.gsc.riken.jp/ Aizawa,K., Akimura,T., Aizawa,T., Carminci,P., Fukuda,S., Hirozane,T., Imocani,K., Ihnii,Y., Itoh,M., Kawai,J., Komno,H., Miyazaki,A., Murata,M., Nakamura,M., Nomura,K., Numasaki,R., Ohno,M., Sakai,K., Sakazume,N., Saesaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M., Waki,K., Wataniki,A., Muranatsu,M. and Hayashizaki,Y. Direct Submission Computational Analysis of Full-length Mouse cDNAs Compared with		

Human Genom. Sequences Mamm. Genome. 12, 673-677 (2001). Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000). RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000). Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001). cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN, Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

SOURCE

```

1. 415
   /organism="Mus musculus"
   /mol_type="mRNA"
   /db_xref="taxon:10090"
   /clone="K630043B16"
   /clone_lib="Riken full-length enriched, pooled tissues
adult B16en, etc."
   /note="(dev_stage=adult,tissue_type=spleen,sex=male),
(dev_stage=adult,tissue_type=kidney,sex=male),
(dev_stage=adult,tissue_type=testis,sex=male),
(dev_stage=adult,tissue_type=hymus,sex=male),
(dev_stage=adult,tissue_type=heart,sex=male),
(dev_stage=adult,tissue_type=colon,sex=male),
(dev_stage=adult,tissue_type=stomach,sex=male),
(dev_stage=adult,tissue_type=livar,sex=male),
(dev_stage=13 days embryo,tissue_type=whole
body,sex=mix),
(dev_stage=14 days embryo,tissue_type=whole
body,sex=mix),
(dev_stage=16 days embryo,tissue_type=whole
body,sex=mix),
(dev_stage=17 days pregnant
body,sex=mix),
(dev_stage=15 days pregnant
adult,tissue_type=anion,sex=female),
(dev_stage=10 day
neonate,tissue_type=drain,sex=mix),
(dev_stage=10 days
neonate,tissue_type=hymus,sex=mix),
(dev_stage=10 days
neonate,tissue_type=heart,sex=mix)"

```

ORIGIN

Query Match 2.38; Score 133.4; DB 5; Length 415;

Matches 137; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 5289 TCCCTGTCAGAACTCTAGATTGTCCTCCAGCTCTGCTACTCTCTCTGCTGTTTC 5348
Db 15 TCCAACTCCAGAACTCTGAGATTGCTCTCCAGCTCTGCTACTCTCTCTCTGCTGTTTC 74

5349 CTCTCTGTCCAGCTGCCCACTGTGGTGCCTCGTTCACAGCTGTGTCCACATTCTCA 5408

Db 75 CTCTCTCCGTCAGCTGCGCCACTGTGGTGCCCTCGTTCACAGCTGTGGTCCACATTCTCA 134

QY 5409 GGATTCTCTGAAGAAGTTAACCAG 5431

Db 135 GGATTCTCTGAAAGTTAACCAG 157

RESULT	13
CBS96657	
LOCUS	
DEFINITION	
CBS96657	739 bp mRNA linear EST 03-APR-2003
AGENCOURT	12990591 NIH_MGC_178 Mus musculus cDNA clone
IMAGE:	30293987 5', mRNA sequence.

SOURCE ORGANISM	Source
Mus musculus (house mouse)	1
Mus musculus	2
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus	3

REFERENCE 1 (bases 1 to 739)
 AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 CDNA Library Preparation: Michael Brownstein Laboratory
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: NDCM64 row: 1 column: 18
 High quality sequence stop: 590.
 Location/Qualifiers

FEATURES
 source
 1..739
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:3029897"
 /lab_host="DH10B (TI-phage-resistant)"
 /clone_lib="NIH_MGC_178"
 /note="Organ: lung and heart; Vector: pDNR-LIB; Site 1: SfiI (ggccatcgcc); Site 2: SfiI (ggccgctggcc); CDNA made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAGGTATACACGAGGCGCCACATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5 kb size fraction. Library created in the laboratory of M. Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 2.3%; Score 133.4; DB 6; Length 739;
 Best Local Similarity 95.8%; Pred. No. 2.1e-23;
 Matches 137; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5289 TCCTGTGAGAGATCTAGATTGGTCTCCAGCCTCTGCTACTCTCTGCTGCTTC 5348
 |||
 Db 17 TCCAACCCAGATCTCTGATGTTGCTCCAGCCTCTGCTACTCTCTGCTGCTTC 76
 QY 5349 CTCCTCTGTCCAGCTGCGCAGCTGTGTGCTGTTCCAGCTGTGTCACATTTCTCA 5408
 |||
 Db 77 CTCCTCTCGTCCAGCTGCGCAGCTGTGTGCTGCTGTTCCAGCTGTGTCACATTTCTCA 136
 QY 5409 GGATTTCTGAAAGTTAACCG 5431
 |||
 Db 137 GGATTTCTGAAAGTTAACCG 159

RESULT 14
 CBS99099 741 bp mRNA linear EST 03-APR-2003
 LOCUS CBS99099
 DEFINITION AGENCOURT_13002811 NIH_MGC_178 Mus musculus CDNA clone
 IMAGE:30298800 5', mRNA sequence.
 ACCESSION CBS99099
 VERSION CBS99099.1 GI:29516955
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 741)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Dr. Michael Brownstein
 CDNA Library Preparation: Michael Brownstein Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)

DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: NDCM61 row: 0 column: 01
 High quality sequence stop: 502.
 Location/Qualifiers

FEATURES
 source
 1..741
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="IMAGE:30298800"
 /lab_host="DH10B (TI-phage-resistant)"
 /clone_lib="NIH_MGC_178"
 /note="Organ: lung and heart; Vector: pDNR-LIB; Site 1: SfiI (ggccatcgcc); Site 2: SfiI (ggccgctggcc); CDNA made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows: 5'-AAGCAGGTATACACGAGGCGCCACATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5 kb size fraction. Library created in the laboratory of M. Brownstein (NIMH, NIH). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 2.3%; Score 133.4; DB 6; Length 741;
 Best Local Similarity 95.8%; Pred. No. 2.1e-23;
 Matches 137; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 5289 TCCTGTGAGAGATCTAGATTGGTCTCCAGCCTCTGCTACTCTCTGCTGCTTC 5348
 |||
 Db 17 TCCAACCCAGATCTCTGATGTTGCTCCAGCCTCTGCTACTCTCTGCTGCTTC 76
 QY 5349 CTCCTCTGTCCAGCTGCGCAGCTGTGTGCTGTTCCAGCTGTGTCACATTTCTCA 5408
 |||
 Db 77 CTCCTCTCGTCCAGCTGCGCAGCTGTGTGCTGCTGTTCCAGCTGTGTCACATTTCTCA 136
 QY 5409 GGATTTCTGAAAGTTAACCG 5431
 |||
 Db 137 GGATTTCTGAAAGTTAACCG 159

RESULT 15
 CA490238 783 bp mRNA linear EST 09-MAR-2004
 LOCUS CA490238
 DEFINITION AGENCOURT_10825991 NIH_MGC_156 Mus musculus CDNA clone
 IMAGE:4755570 5', mRNA sequence.
 ACCESSION CA490238
 VERSION CA490238.1 GI:24953042
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 783)
 NIH-MGC <http://mgs.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Bradfield Laboratory
 CDNA Library Preparation: Mark Blittinger
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
 Plate: LKMO0109 row: 1 column: 01
 High quality sequence stop: 543.
 Location/Qualifiers

FEATURES
 source
 1..783
 /organism="Mus musculus"

This Page Blank (uspto)

Dh 1 GGATCTCGAAGGTCAACAAGGGTCTCCACCACGAGTGCCCTAGTCTCAATTTCAAGT 60
Qy TTCCATGCTTGTCTCACAATGCTGGCTCCCGACAGCTAATTTGGACTTTGTTTAT 120
Dh 61 TTCCATGCTTGTCTCACAATGCTGGCTCCCGACAGCTAATTTGGACTTTGTTTAT 120
Qy 121 TTCAAAAAGGCTGATGAGAGTAGATCTTGTGCTACCCAGCTTAAAGGCTCCCGTGA 180
Dh 121 TTCAAAAAGGCTGATGAGAGTAGATCTTGTGCTACCCAGCTTAAAGGCTCCCGTGA 180
Qy 181 AGCCCTCAGACCTGAGGCTTTTGCACAGCCCTTTAGTGGAGAGCAAAATTAAGCAATTT 240
Dh 181 AGCCCTCAGACCTGAGGCTTTTGCACAGCCCTTTAGTGGAGAGCAAAATTAAGCAATTT 240
Qy 241 TCCTTAAAGCAAAATCTGCTCTAGACTCTTCTTCTGACCTCGGCTCCGAGCTCT 300
Dh 241 TCCTTAAAGCAAAATCTGCTCTCTAGACTCTTCTTCTGACCTCGGCTCCGAGCTCT 300
Qy 301 AGGGTGGGAGGTGGGCTTTGAAAGAAAGTGGGAGTGGCAAAAGCCGATCCCTAG 360
Dh 301 AGGGTGGGAGGTGGGCTTTGAAAGAAAGTGGGAGTGGCAAAAGCCGATCCCTAG 360
Qy 361 GGCCCTGTGAAGTTGGAGGCTTCCCTGTACAGACTGGCTCATATGCTCTCCACAGCC 420
Dh 361 GGCCCTGTGAAGTTGGAGGCTTCCCTGTACAGACTGGCTCATATGCTCTCCACAGCC 420
Qy 421 AAACATAGCAAGAAGTATACCTCTTGTGACTTCCCGAGGCCAGTACCTGTCAAGTT 480
Dh 421 AAACATAGCAAGAAGTATACCTCTTGTGACTTCCCGAGGCCAGTACCTGTCAAGTT 480
Qy 481 GAAACAGAGATTTAGAGAGGCTCTGAACTCACTGAACTGTATCCACCAAGCA 540
Dh 481 GAAACAGAGATTTAGAGAGGCTCTGAACTCACTGAACTGTATCCACCAAGCA 540
Qy 541 AGGACCTAGGTCGCACTGATGATGATCTTACGCTGATTAATGAGAGCTGGGCTAC 600
Dh 541 AGGACCTAGGTCGCACTGATGATGATCTTACGCTGATTAATGAGAGCTGGGCTAC 600
Qy 601 AGAAGTCTGGGGGTGAGGAAGTACCAAGTGAAGTCTTTCAGTGGCAAAAGGTATGACCCC 660
Dh 601 AGAAGTCTGGGGGTGAGGAAGTACCAAGTGAAGTCTTTCAGTGGCAAAAGGTATGACCCC 660
Qy 661 TCAGCAGATGTAATGTCTCCCTTGAATCCCATCCAGGAGGTCTCTAAAGAGACATG 720
Dh 661 TCAGCAGATGTAATGTCTCCCTTGAATCCCATCCAGGAGGTCTCTAAAGAGACATG 720
Qy 721 GGATGAGATGTAATGTGATGGCAATCCAAACACAGTATCCACAGTGTCCCTTGGCCC 780
Dh 721 GGATGAGATGTAATGTGATGGCAATCCAAACACAGTATCCACAGTGTCCCTTGGCCC 780
Qy 781 TTCCACTTAAAGCAGAGAGAGAGAGTAACTTATGCTTCTTCTCCCATCTCCAG 840
Dh 781 TTCCACTTAAAGCAGAGAGAGAGAGTAACTTATGCTTCTTCTCCCATCTCCAG 840
Qy 841 GACACACCCCTGGTCTGCAATATTTCAATTTCTTCTTCAAGTCCCTCTGTGACTTCAT 900
Dh 841 GACACACCCCTGGTCTGCAATATTTCAATTTCTTCTTCAAGTCCCTCTGTGACTTCAT 900
Qy 901 TTGCAAGGCTTTTGAACCTGTCAGCTGTGGAAGATGAGTTTGGGCTTGGTGGCAA 960
Dh 901 TTGCAAGGCTTTTGAACCTGTCAGCTGTGGAAGATGAGTTTGGGCTTGGTGGCAA 960
Qy 961 GCCATCTCAAG 1020
Dh 961 GCCATCTCAAG 1020
Qy 1021 TCATCTGGCGGGCTGGGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Dh 1021 TCATCTGGCGGGCTGGGGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
Qy 1081 TAGTCCCGAGAGATCTGTGCAAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 1140
Dh 1081 TAGTCCCGAGAGATCTGTGCAAGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 1140

Qy 1141 TGAAGAGCAATGGATTTGAGAGACATCTTTGAGTCTCCCTCAACCCCACTACAGAC 1200
Dh 1141 TGAAGAGCAATGGATTTGAGAGACATCTTTGAGTCTCCCTCAACCCCACTACAGAC 1200
Qy 1201 ACACTGTGTGTGGCAGACTCTCTGTTCAACAGCCCTCTGTGTTCTGACCACTGAGCTAG 1260
Dh 1201 ACACTGTGTGTGGCAGACTCTCTGTTCAACAGCCCTCTGTGTTCTGACCACTGAGCTAG 1260
Qy 1261 GCAACAGAGCATGGGCTCTGTGCTGAGGATGAGAGTGGTTACCAATAGCAAAACAG 1320
Dh 1261 GCAACAGAGCATGGGCTCTGTGCTGAGGATGAGAGTGGTTACCAATAGCAAAACAG 1320
Qy 1321 CAGGGAGGAG 1380
Dh 1321 CAGGGAGGAG 1380
Qy 1381 TGCATCTCAGAGAGATGGGAGGCAACACAGCTTTGACACAGAGAGAGAGAGAGAGAGAG 1440
Dh 1381 TGCATCTCAGAGAGATGGGAGGCAACACAGCTTTGACACAGAGAGAGAGAGAGAGAGAG 1440
Qy 1441 AGATTCTGGGCAATTAAGAGAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Dh 1441 AGATTCTGGGCAATTAAGAGAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Qy 1501 ATACCTCATCCGCTCTCCCAATTAAGGCACTCTTCTTCTTATGATCAGACTGAGCTGC 1560
Dh 1501 ATACCTCATCCGCTCTCCCAATTAAGGCACTCTTCTTCTTATGATCAGACTGAGCTGC 1560
Qy 1561 AGCGAAGAGAGCCGTGAGGAGATCACTGTGATGAGAGAGATGTGTGAGAGAGTCCAGG 1620
Dh 1561 AGCGAAGAGAGCCGTGAGGAGATCACTGTGATGAGAGAGATGTGTGAGAGAGTCCAGG 1620
Qy 1621 GCAACCTAAGAGGCAAGGCTTAAAGAGCAAGATTAAGTGTCTTCAAGGTGGCCAGG 1680
Dh 1621 GCAACCTAAGAGGCAAGGCTTAAAGAGCAAGATTAAGTGTCTTCAAGGTGGCCAGG 1680
Qy 1681 CTGTGCAACAAGAGGCTGAGAGACTGTGTGTAGAGCTCAAGATTAAGATGATGCTCAGA 1740
Dh 1681 CTGTGCAACAAGAGGCTGAGAGACTGTGTGTAGAGCTCAAGATTAAGATGATGCTCAGA 1740
Qy 1741 ATGGGCGGGGGGGGAGATTTCTGGGGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Dh 1741 ATGGGCGGGGGGGGAGATTTCTGGGGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800
Qy 1801 AGAATCTGGAAGGCTGGAAGACATACATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Dh 1801 AGAATCTGGAAGGCTGGAAGACATACATTAAGGAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
Qy 1861 TAAATCATGAAG 1920
Dh 1861 TAAATCATGAAG 1920
Qy 1921 CATGGAAG 1980
Dh 1921 CATGGAAG 1980
Qy 1981 CCTGTGTATGAG 2040
Dh 1981 CCTGTGTATGAG 2040
Qy 2041 GTAGAGACAG 2100
Dh 2041 GTAGAGACAG 2100
Qy 2101 CTGTGGCAGAGAGGTCATCTGTGTCAACATCTCTGAGATGTTCAATCAAGCTGACGGGC 2160
Dh 2101 CTGTGGCAGAGAGGTCATCTGTGTCAACATCTCTGAGATGTTCAATCAAGCTGACGGGC 2160
Qy 2161 TGCTTTGGAGAGAGCTGGAAGAGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
Dh 2161 TGCTTTGGAGAGAGCTGGAAGAGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220

2221 AAGGAAGAGGGCAAAACGAGCCACACAAAGAGGAGAGCCAGAACTGATTAATC 2280
2222 AAGGAAGAGGGCAAAACGAGCCACACAAAGAGGAGAGCCAGAACTGATTAATC 2280
2281 CTTCCTTGTGATCTTCCATAGAGGAGAGTGGAACTCTGTGACCAACATCCCATGA 2340
2282 CTTCCTTGTGATCTTCCATAGAGGAGAGTGGAACTCTGTGACCAACATCCCATGA 2340
2341 GGGCCCACTAACCCATACCAAGTTTGGCTGAGTGGATTTAGTTCCTTGAGGAGAG 2400
2342 GGGCCCACTAACCCATACCAAGTTTGGCTGAGTGGATTTAGTTCCTTGAGGAGAG 2400
2401 CCTGGCTTGTCTCTTGGACCTGACCCCAAGCTGACCCCAATGTTCTCAGTACCTTGAAT 2460
2402 CCTGGCTTGTCTCTTGGACCTGACCCCAAGCTGACCCCAATGTTCTCAGTACCTTGAAT 2460
2461 GGGCTCAAGAGCTTGAAGAACGAGGAGTGCATATTTAGGCTATGAGGCTTAACTGAGGCT 2520
2462 GGGCTCAAGAGCTTGAAGAACGAGGAGTGCATATTTAGGCTATGAGGCTTAACTGAGGCT 2520
2521 TGGACACAGAGGCTCAAGTGAACCTCAGGAGACAGCTGACAGAGTGGCTTTATCC 2580
2522 TGGACACAGAGGCTCAAGTGAACCTCAGGAGACAGCTGACAGAGTGGCTTTATCC 2580
2581 CCNAAGAGCAACCATTTGGCATAGTGGCTGCAATAGGGAATGCAAGTTGAATCAGTTC 2640
2582 CCNAAGAGCAACCATTTGGCATAGTGGCTGCAATAGGGAATGCAAGTTGAATCAGTTC 2640
2583 CCNAAGAGCAACCATTTGGCATAGTGGCTGCAATAGGGAATGCAAGTTGAATCAGTTC 2640
2641 CCTTCAAGATATCTGATGCAAGACCTTAAAGCCCTGAGAGAGGGGTATGCTCTGCCC 2700
2642 CCTTCAAGATATCTGATGCAAGACCTTAAAGCCCTGAGAGAGGGGTATGCTCTGCCC 2700
2701 CCACCCACCATTAAGGGAGAGTGAATCTATCCTAGAGGGGCTGGCCACTTGGGAGAGACAC 2760
2702 CCACCCACCATTAAGGGAGAGTGAATCTATCCTAGAGGGGCTGGCCACTTGGGAGAGACAC 2760
2761 ATTACTGAGAGTGTGAGCCAGAGAAAACTGACCGGCTGTGCTGCTGACCACTCCACAC 2820
2762 ATTACTGAGAGTGTGAGCCAGAGAAAACTGACCGGCTGTGCTGCTGACCACTCCACAC 2820
2821 TCTAGAGCTATATTGAGAGGTGACAGTGAATAGGAGTGGAGCTGTGAGCAGGAGAGTGT 2880
2822 TCTAGAGCTATATTGAGAGGTGACAGTGAATAGGAGTGGAGCTGTGAGCAGGAGAGTGT 2880
2881 TCTGAGGTGTGAGAGGTGTGAGGGAGAAAGCCAGAGCAGGAGTGTGCTTGTCTCTGAA 2940
2882 TCTGAGGTGTGAGAGGTGTGAGGGAGAAAGCCAGAGCAGGAGTGTGCTTGTCTCTGAA 2940
2941 CACAATGTCTACTTATTTAAACAGGAGTGAACCTGCTAAAGAACCAACATCTACGACCTC 3000
2942 CACAATGTCTACTTATTTAAACAGGAGTGAACCTGCTAAAGAACCAACATCTACGACCTC 3000
3001 TGAAGAGACAGCAGCCCTGAGAGACAGGGGTGTCTCTGAGCCTTGGGTGATGAGTGTG 3060
3002 TGAAGAGACAGCAGCCCTGAGAGACAGGGGTGTCTCTGAGCCTTGGGTGATGAGTGTG 3060
3061 CCACAAAGAGGAGCAGTGTGATGATTAAGGCCCCAGAGCGCTTGAAGAGGAGGCACTTG 3120
3062 CCACAAAGAGGAGCAGTGTGATGATTAAGGCCCCAGAGCGCTTGAAGAGGAGGCACTTG 3120
3121 GGAAGAGGTCAGTGTGAGAGCCCTATCATGATGAATCTGAGCTGTGGGCCCAACTGTGTG 3180
3122 GGAAGAGGTCAGTGTGAGAGCCCTATCATGATGAATCTGAGCTGTGGGCCCAACTGTGTG 3180
3181 TAAATCTTGTGGGCTGTCCAGGCAATTAAGAGCAGACCTGTGACCTGTGGAGCTGGAGA 3240
3182 TAAATCTTGTGGGCTGTCCAGGCAATTAAGAGCAGACCTGTGACCTGTGGAGCTGGAGA 3240
3241 GGGGGAAGGAGAGAACCCCACTTATACCTTTCCTCAGGCCCCAGAGATTTAAACCT 3300
3242 GGGGGAAGGAGAGAACCCCACTTATACCTTTCCTCAGGCCCCAGAGATTTAAACCT 3300
3301 CTGGCTTCCCTTCCACCTTCCATCAAGAGTGAAGGGTGTGCAAGGAGGAGGTAAAAA 3360

3302 CTGGCTTCCCTTCCACCTTCCATCAAGAGTGAAGGGTGTGCAAGGAGGAGGTAAAAA 3360
3361 CCTACATGTCCAAACATCATGTGTGACAGATATATGATCATATGTGTGAGGCAAGAA 3420
3362 CCTACATGTCCAAACATCATGTGTGACAGATATATGATCATATGTGTGAGGCAAGAA 3420
3421 GGAAGATCTGAGGCTTAACTGAGGTTAATGTGTAAGTCTGTGATGTGTGTGTG 3480
3422 GGAAGATCTGAGGCTTAACTGAGGTTAATGTGTAAGTCTGTGATGTGTGTGTG 3480
3481 ACTGAAACGAGGACATGCTGTGACAGCTGTTCAGTTCTGTGCGTGAAGTTACCAAGCTGCA 3540
3482 ACTGAAACGAGGACATGCTGTGACAGCTGTTCAGTTCTGTGCGTGAAGTTACCAAGCTGCA 3540
3541 GGTGTGTGTGTAATTTGGCCAAAGGCAAGTGGGTAAATCCTTCCATGCTTTAAAGAGAT 3600
3542 GGTGTGTGTGTAATTTGGCCAAAGGCAAGTGGGTAAATCCTTCCATGCTTTAAAGAGAT 3600
3601 TGGATGATGGGCTGATCTCAAGAACATGGAATTAAGATGGAACCTATATGTGTCT 3660
3602 TGGATGATGGGCTGATCTCAAGAACATGGAATTAAGATGGAACCTATATGTGTCT 3660
3661 CTAACTAAGGTAGCAAGGCTTTTGGAGGACACTGTCTAGAGATGTGGGCAACAGAGAC 3720
3662 CTAACTAAGGTAGCAAGGCTTTTGGAGGACACTGTCTAGAGATGTGGGCAACAGAGAC 3720
3721 TACAGACATATCTGTACAGATTAAGAGAGAGAGAGGGGCTGTAAATCTCTTACTA 3780
3722 TACAGACATATCTGTACAGATTAAGAGAGAGAGAGAGGGGCTGTAAATCTCTTACTA 3780
3781 TGAAGGGAACATGAGTGTGACCTGCAAAATGTGATGCTCCCTAGACATCATGACTT 3840
3782 TGAAGGGAACATGAGTGTGACCTGCAAAATGTGATGCTCCCTAGACATCATGACTT 3840
3841 TGTCTCTGGGAGCCAGACACTGTGAACTTCAAGTCTGAGAGATGAGAGGCTCCCTCA 3900
3842 TGTCTCTGGGAGCCAGACACTGTGAACTTCAAGTCTGAGAGATGAGAGGCTCCCTCA 3900
3901 GCTTGAAGTATGACAGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3960
3902 GCTTGAAGTATGACAGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3960
3961 TGTGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4020
3962 TGTGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4020
4021 TGTTCACACTCACCTTCAGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4080
4022 TGTTCACACTCACCTTCAGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4080
4081 GTGTGAGAGCTCTGTGTCTCTCTGCAATGCTGAGGCTTGTGGGAGGAGGAGGAGGAGG 4140
4082 GTGTGAGAGCTCTGTGTCTCTCTGCAATGCTGAGGCTTGTGGGAGGAGGAGGAGGAGG 4140
4141 GGGCAAGGACATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4200
4142 GGGCAAGGACATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4200
4201 GGAAGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4260
4202 GGAAGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4260
4261 TTCTAACGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4320
4262 TTCTAACGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4320
4321 AAGTGAAGTTTACCACTCCATCATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4380
4322 AAGTGAAGTTTACCACTCCATCATGATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4380
4381 CCTATCAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4440

Db 4381 CCTATCATGATATGAGAAAAAGTGAAGTTCAGTTCCTCATCTAGTATAGAA 4440
Qy 4441 AAGTGAAGTCGATTTACCACTCCCTATCATGATAGAGAAAAAGTGAAGTTC 4500
Db 4441 AAGTGAAGTCGATTTACCACTCCCTATCATGATAGAGAAAAAGTGAAGTTC 4500
Qy 4501 CCACTCCCTATCATGATAGAGAAAAAGTGAAGTTCAGTTCCTCATCTAGTAT 4560
Db 4501 CCACTCCCTATCATGATAGAGAAAAAGTGAAGTTCAGTTCCTCATCTAGTAT 4560
Qy 4561 AGAGAAAGTGAAGTTCAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4620
Db 4561 AGAGAAAGTGAAGTTCAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4620
Qy 4621 TATGAGATGAGATATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4680
Db 4621 TATGAGATGAGATATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4680
Qy 4681 GGTATGAGAGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4740
Db 4681 GGTATGAGAGAGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4740
Qy 4741 GGAAGTCCCTTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4800
Db 4741 GGAAGTCCCTTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4800
Qy 4801 CCAAGAGAGAGAGATTCAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4860
Db 4801 CCAAGAGAGAGAGATTCAGTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4860
Qy 4861 TCCCTTACGCTTCCTTCATCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 4920
Db 4861 TCCCTTACGCTTCCTTCATCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 4920
Qy 4921 TTGGCTTCCTTCCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4980
Db 4921 TTGGCTTCCTTCCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4980
Qy 4981 CTCCTGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 5040
Db 4981 CTCCTGCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 5040
Qy 5041 TTTTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 5100
Db 5041 TTTTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 5100
Qy 5101 CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 5160
Db 5101 CTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC 5160
Qy 5161 CTGCTGTCAGAGTCGAGAGATTCACACCTGAGAGATTCACACCTGAGAGATTCAC 5220
Db 5161 CTGCTGTCAGAGTCGAGAGATTCACACCTGAGAGATTCACACCTGAGAGATTCAC 5220
Qy 5221 GTGAGCACAAGCTTCAGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5280
Db 5221 GTGAGCACAAGCTTCAGTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5280
Qy 5281 CTGCTGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5340
Db 5281 CTGCTGTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5340
Qy 5341 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5400
Db 5341 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5400
Qy 5401 ATTCTTACAGATTCCTGAG 5460
Db 5401 ATTCTTACAGATTCCTGAG 5460
Qy 5461 TCACGATTCCTCCGAG 5520
Db 5461 TCACGATTCCTCCGAG 5520

Qy 5521 CTCTTAGCAAACTCAGGACCTTACCCACATAGACCTTGAAGAGAGAGAGAGAGAGAG 5580
Db 5521 CTCTTAGCAAACTCAGGACCTTACCCACATAGACCTTGAAGAGAGAGAGAGAGAGAG 5580
Qy 5581 TTACATGAGAGTCCTGCTGAG 5640
Db 5581 TTACATGAGAGTCCTGCTGAG 5640
Qy 5641 GGTATGAGAGAGTCAGAGATTCATAG 5700
Db 5641 GGTATGAGAGAGTCAGAGATTCATAG 5700
Qy 5701 TCCCTCTATCTCCCTCATAGAGATTCAG 5760
Db 5701 TCCCTCTATCTCCCTCATAGAGATTCAG 5760

RESULT 2
US-10-024-066-5
; Sequence 5, Application US/10024066
; Publication No. US2002016134A1
; GENERAL INFORMATION:
; APPLICANT: Field, Loren J.
; APPLICANT: Pasmartch, Kishore Babu S.
; TITLE OF INVENTION: CARDIOMYOCYTES WITH ENHANCED PROLIFERATIVE POTENTIAL,
; FILE REFERENCE: 7037-450
; CURRENT APPLICATION NUMBER: US/10/024, 066
; PRIOR FILING DATE: 2001-12-18
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: PCT/US00/16827
; PRIOR FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 5443
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-024-066-5

Query Match 88.0%; Score 5045; DB 13; Length 5443;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 5383; Conservative 0; Mismatches 60; Indels 292; Gaps 1;

Qy 1 GGAATCCGCAAGGTCAACAAGGCTCCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Db 1 GGAATCCGCAAGGTCAACAAGGCTCCACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
Qy 61 TTCCATGCTTGTTCACATGCTGCTCCCAAGCTTAATTTGATTTTAT 120
Db 61 TTCCATGCTTGTTCACATGCTGCTCCCAAGCTTAATTTGATTTTAT 120
Qy 121 TTAAAAAGGCTGAGATGAGAGATGATTTGCTTACCAAGCTCTTAAGAGAGAGAGAGAGAGAG 180
Db 121 TTAAAAAGGCTGAGATGAGAGATGATTTGCTTACCAAGCTCTTAAGAGAGAGAGAGAGAGAG 180
Qy 181 AGCCCTCAGAGCTGAGAGCTTTTCAACAGAGCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Db 181 AGCCCTCAGAGCTGAGAGCTTTTCAACAGAGCTTTTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
Qy 241 TCCCTTAAGCAAAATCTGCTCTAGAGCTTTCTTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Db 241 TCCCTTAAGCAAAATCTGCTCTAGAGCTTTCTTCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
Qy 301 AGGCTGAGAGAGCTGAGAGCTTTGAG 360
Db 301 AGGCTGAGAGAGCTGAGAGCTTTGAG 360
Qy 361 GGGCTGTGAGAGTTCGAGAGCTTCCTGTAAG 420
Db 361 GGGCTGTGAGAGTTCGAGAGCTTCCTGTAAG 420

Qy	421	AAACATGCAAGAAAGTGAATACCTCTTGTGTGACTCTCCAGGCCAGTACTGTGAGTT	480
Dp	421	AAACATGCAAGAAAGTGAATACCTCTTGTGTGACTCTCCAGGCCAGTACTGTGAGTT	480
Qy	481	GAAACAGGATTTAGAGAAAGCTCTGAACTCACTGAACTCTGAAAGTCAATCAACAAGCA	540
Dp	481	GAAACAGGATTTAGAGAAAGCTCTGAACTCACTGAACTCTGAAAGTCAATCAACAAGCA	540
Qy	541	AGCACTTAGGTGCCACTGTAGTTAGTATCCTAGCGTGATATAATGACAGCTGGGCAC	600
Dp	541	AGCACTTAGGTGCCACTGTAGTTAGTATCCTAGCGTGATATAATGACAGCTGGGCAC	600
Qy	601	AGAAGTCTCGGGGTGTAGGAACTGACACAGTGCATTTTCAAGTCGGCAAAAGTATACCCCC	660
Dp	601	AGAAGTCTCGGGGTGTAGGAACTGACACAGTGCATTTTCAAGTCGGCAAAAGTATACCCCC	660
Qy	661	TCAGCAGATGATGATATGTCCCTTAGATCCCATCCAGGAGAGTCTCTAAGAGGACATG	720
Dp	661	TCAGCAGATGATGATATGTCCCTTAGATCCCATCCAGGAGAGTCTCTAAGAGGACATG	720
Qy	721	GGATGAGAGATGTGTGTCAATGTGGCAATTCAAACACAGTATCCACAGTGTCCCTTGCCCC	780
Dp	721	GGATGAGAGATGTGTGTCAATGTGGCAATTCAAACACAGTATCCACAGTGTCCCTTGCCCC	780
Qy	781	TTTCCACTTAGCAGAGAGACAGTAACTTTAGCCTATCTTTCCTGCCATCCGCCAC	840
Dp	781	TTTCCACTTAGCAGAGAGACAGTAACTTTAGCCTATCTTTCCTGCCATCCGCCAC	840
Qy	841	GACACACCCCTGTGTCTGACAGTATTCATTTCTTCCCTGACGTCCCTGTGTACTTCAT	900
Dp	841	GACACACCCCTGTGTCTGACAGTATTCATTTCTTCCCTGACGTCCCTGTGTACTTCAT	900
Qy	901	TTTGAAGGCTTTTACCTGTGACGTGTGTGAAGATATAGTTTGGCCCTAGGTGTGGCAA	960
Dp	901	TTTGAAGGCTTTTACCTGTGACGTGTGTGAAGATATAGTTTGGCCCTAGGTGTGGCAA	960
Qy	961	GCCATCTTCAAGAGAAAGCAGACAACAGGGGACAGATTTTGGAAAGATTCAGAACTTAA	1020
Dp	961	GCCATCTTCAAGAGAAAGCAGACAACAGGGGACAGATTTTGGAAAGATTCAGAACTTAA	1020
Qy	1021	TCACTGGCGGGCTGTGGGGTGTGAAAAAAGATGATGTGTCCGCTCCAGCTAAGCAAGC	1080
Dp	1021	TCACTGGCGGGCTGTGGGGTGTGAAAAAAGATGATGTGTCCGCTCCAGCTAAGCAAGC	1080
Qy	1081	TAGTCTCCCGAATATCTGTGCAACAGCTGGGTGTGCTGGGGGTATAGCTTATGGAATGTGGTCT	1140
Dp	1081	TAGTCTCCCGAATATCTGTGCAACAGCTGGGTGTGCTGGGGGTATAGCTTATGGAATGTGGTCT	1140
Qy	1141	TGAAGAACAATGGGATTTGGAAGACATCTCTTTGAGTCTCCCTCAACCCCACTTACAGAC	1200
Dp	1141	TGAAGAACAATGGGATTTGGAAGACATCTCTTTGAGTCTCCCTCAACCCCACTTACAGAC	1200
Qy	1201	ACACTCGGTGTGGCCAGACTCTCTGTTCACAGCCCTCTGTGTCTGACCACTAGAGCTAG	1260
Dp	1201	ACACTCGGTGTGGCCAGACTCTCTGTTCACAGCCCTCTGTGTCTGACCACTAGAGCTAG	1260
Qy	1261	GCAACCAAGAGCATGGGCTCTGTGCTGTGAGAGTGAAGAGTTGTGTTTCAATAGCAAAAAACAG	1320
Dp	1261	GCAACCAAGAGCATGGGCTCTGTGCTGTGAGAGTGAAGAGTTGTGTTTCAATAGCAAAAAACAG	1320
Qy	1321	CAGGGAGAGGAGAAACAAGAACGAATATAGAGAGAGAGAAAGGCCCACTCAATCAGA	1380
Dp	1321	CAGGGAGAGGAGAAACAAGAACGAATATAGAGAGAGAGAAAGGCCCACTCAATCAGA	1380
Qy	1381	TGCAAGTCAAGAGAGATGGGAGCCCAACAAGCTTTGAGCAGAGAGAAACAAGAAAGGGAG	1440
Dp	1381	TGCAAGTCAAGAGAGATGGGAGCCCAACAAGCTTTGAGCAGAGAGAAACAAGAAAGGGAG	1440
Qy	1441	AGATTCTTGGGCAATAAGAGGCCACAAGAAAGAGGCCAGGCCCTCCCAAGTCTCTCTTT	1500
Dp	1441	AGATTCTTGGGCAATAAGAGGCCACAAGAAAGAGGCCAGGCCCTCCCAAGTCTCTCTTT	1500

QY	1501	ATACCTCATCCGCTCTCCCAATTAAAGCCCACTTTCTTCTTAATACAGACTTGACGTCC	1500
Db	1501	ATACCCTCATCCCGCTCTCCCAATTAAAGCCCACTTTCTTCTTAATACAGACTTGACGTCC	1500
QY	1561	AGCGAAGGACCCGCTAGGGAGGATCAACCTGGAATGAAAGGAGATGTGTGAGAAAGCTCCAGG	1620
Db	1561	AGCGAAGGACCCGCTAGGGAGGATCAACCTGGAATGAAAGGAGATGTGTGAGAAAGCTCCAGG	1620
QY	1621	GCAACCTTAAGAGCCGAGAGCCCTTAAGAGCAAGAGATTAAGGTGCTTCAAAAGTGGCCAGG	1680
Db	1621	GCAACCTTAAGAGCCGAGAGCCCTTAAGAGCAAGAGATTAAGGTGCTTCAAAAGTGGCCAGG	1680
QY	1681	CTGTGACACACAGAGGGTGCAGAGACTGTGGTAGACCTCAAGATTAAGATGATGCTCAGA	1740
Db	1681	CTGTGACACACAGAGGGTGCAGAGACTGTGGTAGACCTCAAGATTAAGATGATGCTCAGA	1740
QY	1741	ATGGGCGGGGGGGGGGATTTCTGGGGGGGGGGAGAGAAAGGTGAAGAGAGCTGGAACAG	1800
Db	1741	ATGGGCGGGGGGGGGGATTTCTGGGGGGGGGGAGAGAAAGGTGAAGAGAGCTGGAACAG	1800
QY	1801	AGAACTCTGGAAGCCGCTGGAAAACGATTACATTAAGGGAAGAACCCGAGCTACTTTAGATG	1860
Db	1801	AGAACTCTGGAAGCCGCTGGAAAACGATTACATTAAGGGAAGAACCCGAGCTACTTTAGATG	1860
QY	1861	TAAATCATGAAAAGACAGGGAGAGGGAGAGCTGAGAGAGTGAAGAGACCCCGGGGCAAGA	1920
Db	1861	TAAATCATGAAAAGACAGGGAGAGGGAGAGCTGAGAGAGTGAAGAGACCCCGGGGCAAGA	1920
QY	1921	CATGGAAGCAAGGACCAAGCCGATTTGAGCCGCTCCGTAAATCACCCTGCTGAAGGCAAG	1980
Db	1921	CATGGAAGCAAGGACCAAGCCGATTTGAGCCGCTCCGTAAATCACCCTGCTGAAGGCAAG	1980
QY	1981	CCCTGTATATGAGCACCAAGAACAGAGAGGCTAAGGGTTAATGTCCAGACAGGGAAACAGAG	2040
Db	1981	CCCTGTATATGAGCACCAAGAACAGAGAGGCTAAGGGTTAATGTCCAGACAGGGAAACAGAG	2040
QY	2041	GTAGAACAAGGAACAGACAGAGACGGGGAGCCAGGTAAACAAGAAATGTGCTTCTTCAAC	2100
Db	2041	GTAGAACAAGGAACAGACAGAGACGGGGAGCCAGGTAAACAAGAAATGTGCTTCTTCAAC	2100
QY	2101	CTGTGACCAAGAGCGTCCATCTGTGTCCAATATCTTGAATGTTCAATCAAGCTGCAAGGAC	2160
Db	2101	CTGTGACCAAGAGCGTCCATCTGTGTCCAATATCTTGAATGTTCAATCAAGCTGCAAGGAC	2160
QY	2161	TGGCTGTGGGAAGGACGCTGGAAGAGTATGTGAGAGCCAGGGGAGACAAGGGGGGCTTGA	2220
Db	2161	TGGCTGTGGGAAGGACGCTGGAAGAGTATGTGAGAGCCAGGGGAGACAAGGGGGGCTTGA	2220
QY	2221	AAGGAAGAAGAGGGCCAACCAAGGCCACACAAGAGGGGAGAGGCCCAAGACTGATGTTAACTC	2280
Db	2221	AAGGAAGAAGAGGGCCAACCAAGGCCACACAAGAGGGGAGAGGCCCAAGACTGATGTTAACTC	2280
QY	2281	CTTCTCTGTGTCATCTTCCATAGAGAGGAGTGGGAATCTGTGTACCAACATCCCCATGA	2340
Db	2281	CTTCTCTGTGTCATCTTCCATAGAGAGGAGTGGGAATCTGTGTACCAACATCCCCATGA	2340
QY	2341	GCCCCCACTACCAATCCAAAGTTTGGGCTGAGTGTGGCAATTCAGAGTCCCTGAGAGCAGAG	2400
Db	2341	GCCCCCACTACCAATCCAAAGTTTGGGCTGAGTGTGGCAATTCAGAGTCCCTGAGAGCAGAG	2400
QY	2401	CCTGAGCCTTGTCTCTTGGACCTGACCCCAAGCTGACCCCAATGTCTCAGTACCTTTGAAT	2460
Db	2401	CCTGAGCCTTGTCTCTTGGACCTGACCCCAAGCTGACCCCAATGTCTCAGTACCTTTGAAT	2460
QY	2461	GCCCTCAAGAGCTTGAAGAACCAAGGCAAGTACATTTAGCCCATGAGGCTTAACCCCTGAGCT	2520
Db	2461	GCCCTCAAGAGCTTGAAGAACCAAGGCAAGTACATTTAGCCCATGAGGCTTAACCCCTGAGCT	2520
QY	2521	TGCAACAAGAGGCTCAAGATGACTTCAAGGGAGCAACGTGCAGACAGGTGGCTTTATTC	2580
Db	2521	TGCAACAAGAGGCTCAAGATGACTTCAAGGGAGCAACGTGCAGACAGGTGGCTTTATTC	2580
QY	2581	CCAAAGAGCAACCAATTGGCAATAGGTGGCTGCAANTGGATGCAATGCAAGTTGAATCAGGTCC	2640

Db	2581	CCAAAGGCAACCAATTGGCATAGTGGCTGCGCAATGCGAAATGGCAAGGTTGAATCAAGTCC	2640
Qy	2641	CCCTTCAAGAAATACATGCAATGCAAGACCTTAAGACCCCTGGAAGAGAGGGGTATATGCTCTGCC	2700
Db	2641	CCTTCAGAAATATCTGATGCAAGACCTTAAGACCCCTGGAAGAGAGGGGTATATGCTCTGCC	2700
Qy	2701	CCACCCACCATTAAGGGGAGTGAACCTATCTTAAGGGGGCTGGCGACCTTTGGGGAGACACAC	2760
Db	2701	CCACCCACCATTAAGGGGAGTGAACCTATCTTAAGGGGGCTGGCGACCTTTGGGGAGACACAC	2760
Qy	2761	ATTACTGAGATGCTGAGGCCCGCAGAAAACTGACCGCCCTGATCTCTCCACCTCCACAC	2820
Db	2761	ATTACTGAGATGCTGAGGCCCGCAGAAAACTGACCGCCCTGATCTCTCCACCTCCACAC	2820
Qy	2821	TCTAGAGCTTAATTAGAGAGTGAAGTGAATAGGGTGGGAGCTGTAGCAGAGGAGATGT	2880
Db	2821	TCTAGAGCTTAATTAGAGAGTGAAGTGAATAGGGTGGGAGCTGTAGCAGAGGAGATGT	2880
Qy	2881	TCTGGGTGTGAGGGGTGTAGGGGAAAGCCAGAGCAGGGAGTCTGGCTTTGTCTCTGAA	2940
Db	2881	TCTGGGTGTGAGGGGTGTAGGGGAAAGCCAGAGCAGGGAGTCTGGCTTTGTCTCTGAA	2940
Qy	2941	CACATGTCTACTTAATATTAACAGGCACTGACCTGTAAACCCCAATCTAAGACCTC	3000
Db	2941	CACATGTCTACTTAATATTAACAGGCACTGCTAAACCCCAATCTAAGACCTC	3000
Qy	3001	TGAAAGAAGACAGAGCCCTGGAAGACAGGGGTTGTCTGAGCCCTTGGGTCTTATGATG	3060
Db	3001	TGAAAGAAGACAGAGCCCTGGAAGACAGGGGTTGTCTGAGCCCTTGGGTCTTATGATG	3060
Qy	3061	CCACAAAGAGAGGCACTGAATGTGAATTAAGGCCCCAGAGAGCTTTAGAAAGGCACTTG	3120
Db	3061	CCACAAAGAGAGGCACTGAATGTGAATTAAGGCCCCAGAGAGCTTTAGAAAGGCACTTG	3120
Qy	3121	GGAAGGGGTCAATCTGCGAGAGCCCTTATCCATGGAATTTGAGGCTTGGGGCCACTGGTG	3180
Db	3121	GGAAGGGGTCAATCTGCGAGAGCCCTTATCCATGGAATTTGAGGCTTGGGGCCACTGGTG	3180
Qy	3181	TAAATCTCTGGGGCTGGCAGGCACTTCAAGAGCACTTGACATCCCTGAGCCAGCCGGGGA	3240
Db	3181	TAAATCTCTGGGGCTGGCAGGCACTTCAAGAGCACTTGACATCCCTGAGCCAGCCGGGGA	3240
Qy	3241	GCGGGAAGGAGCAACCCGCCCACTTATACCTTTCTCCAGGCCAGATTAACACT	3300
Db	3241	GCGGGAAGGAGCAACCCGCCCACTTATACCTTTCTCCAGGCCAGATTAACACT	3300
Qy	3301	CTGGCCTTCCCTTCCCACTCCCATGAGAGTGAAGGGTTGCGAGAGGGGTAAAAA	3360
Db	3301	CTGGCCTTCCCTTCCCACTCCCATGAGAGTGAAGGGTTGCGAGAGGGGTAAAAA	3360
Qy	3361	CCTACATGTCCAAACATCATGGTGTGACATATATAGATCAGTATGTGAGAGGCAAGAA	3420
Db	3361	CCTACATGTCCAAACATCATGGTGTGACATATATAGATCAGTATGTGAGAGGCAAGAA	3420
Qy	3421	GGAATCTGCAAGGCTTAACTGGGTTAATGTGTAAGCTGTGTGATGTGTGTCTG	3480
Db	3421	GGAATCTGCAAGGCTTAACTGGGTTAATGTGTAAGCTGTGTGATGTGTGTCTG	3480
Qy	3481	ACTGAAAAACGGGCATGCTGTGCACTGTTCAAGTTCTGTGGGTGAGGTTAACAGACTGCA	3540
Db	3481	ACTGAAAAACGGGCATGCTGTGCACTGTTCAAGTTCTGTGGGTGAGGTTAACAGACTGCA	3540
Qy	3541	GCTTGTGTGTAAATTTGCCAAGGGAAGTGGGTGAATCCCTTCATGCTTTAAAGAT	3600
Db	3541	GCTTGTGTGTAAATTTGCCAAGGGAAGTGGGTGAATCCCTTCATGCTTTAAAGAT	3600
Qy	3601	TGATATGATGACCTGCATCTCAAGACCAATGAAAAATAGATGCACTCTATATGTGCT	3660
Db	3601	TGATATGATGACCTGCATCTCAAGACCAATGAAAAATAGATGCACTCTATATGTGCT	3660
Qy	3661	CTAAGCTAAGTGAAGGCTCTTTGAGAGCACTGTCTAGATGTGGGCAACAGAGAC	3720

Db	3661	CTTAAGCTAAGGTAGCAAAAGGCTTTTGGAGAGACACTGTCTTAGAATGTGGCCAAACAGAGAC	3720
Qy	3721	TACAGACAGTATCTGTATCAGAGTAAGAGAGAGAGAGGGGGGTGTAGAAATTTCTTTACTA	3780
Db	3721	TACAGACAGTATCTGTATCAGAGTAAAGAGAGAGAGGGGGGTGTAGAAATTTCTTTACTA	3780
Qy	3781	TCAAAAGGAAAATCTAAGTCGTGACACTCTGCAAAGTGGATGCTCTCCCTAGACATCATGTGACTT	3840
Db	3781	TCAAAAGGAAAATCTAAGTCGTGACACTCTGCAAAGTGGATGCTCTCCCTAGACATCATGTGACTT	3840
Qy	3841	TGTCTCTGGGAGGCACACACTGTGGAATCTCAGGCTCTGAGAGAGTAGAGAGCTCCCTCA	3900
Db	3841	TGTCTCTGGGAGGCACACACTGTGGAATCTCAGGCTCTGAGAGAGTAGAGAGCTCCCTCA	3900
Qy	3901	GCCTGAAGCTATGCGAGTATGCCAGGGTTGAAAGGGGGGAAAGGAGAGCCTTGGGATGGAGAC	3960
Db	3901	GCCTGAAGCTATGCGAGTATGCCAGGGTTGAAAGGGGGGAAAGGAGAGCCTTGGGATGGAGAC	3960
Qy	3961	TTGTGTGTGGAGGACAGGGGACAAATATTAAGCTTGGAAAGAGAGTGAACCTTTAACCCAG	4020
Db	3961	TTGTGTGTGGAGGACAGGGGACAAATATTAAGCTTGGAAAGAGAGTGAACCTTTAACCCAG	4020
Qy	4021	TTGTTCAACTCAGCCCTTCAGATTTAAATAATACAGGTAAGGGGCTGGGTAGGGGAGGTG	4080
Db	4021	TTGTTCAACTCAGCCCTTCAGATTTAAATAATACAGGTAAGGGGCTGGGTAGGGGAGGTG	4080
Qy	4081	GTGTGAGACGCTCCTGTCTCTCTCTGATGATGCCCTGAGGCCCTTTTGGGGAGAGAAATGT	4140
Db	4081	GTGTGAGACGCTCCTGTCTCTCTCTGATGATGCCCTGAGGCCCTTTTGGGGAGAGAAATGT	4140
Qy	4141	GCCCAAGGACTTAAAAAAAAGGCCATGAGACCAGAGGGGCGAGGGCAACAGACTTTCAATGG	4200
Db	4141	GCCCAAGGACTTAAAAAAAAGGCCATGAGACCAGAGGGGCGAGGGCAACAGACTTTCAATGG	4200
Qy	4201	GCAAAACCTTGGGGGCCCGTAAAGATGATGATGATCAAGAACTGGCCAAATGATATCCCTTCTTC	4260
Db	4201	GCAAAACCTTGGGGGCCCGTAAAGATGATGATGATGATCAAGAACTGGCCAAATGATATCCCTTCTTC	4260
Qy	4261	TTCTTACCGGACAGAGAGGAACTCGAGTTTACCACTCCCTATCAGTGAATAGAAAAATGA	4320
Db	4261	TTCTTACCGGACAGAGAGGAACTCGAGTTTACCACTCCCTATCAGTGAATAGAAAAATGA	4320
Qy	4321	AAATCGAGTTTACCACTCCCTATCAGTGAATGAGAAAAATGAAAGTCGAGTTTACCACTC	4380
Db	4321	AAATCGAGTTTACCACTCCCTATCAGTGAATGAGAAAAATGAAAGTCGAGTTTACCACTC	4380
Qy	4381	CCTATCAGTGAATAGAGAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAA	4440
Db	4381	CCTATCAGTGAATAGAGAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAA	4440
Qy	4441	AAATGAAAGTCGAGTTTACCACTCCCTATCAGTGAATAGAGAAAAATGAAAGTCGAGTTTA	4500
Db	4441	AAATGAAAGTCGAGTTTACCACTCCCTATCAGTGAATAGAGAAAAATGAAAGTCGAGTTTA	4500
Qy	4501	CCACTCCTTATCAGTGAATAGAGAAAAATGAAAGTCGAGTTTACCACTCCCTATCAGTAT	4560
Db	4501	CCACTCCTTATCAGTGAATAGAGAAAAATGAAAGTCGAGTTTACCACTCCCTATCAGTAT	4560
Qy	4561	AGAGAAAAGTAAAGTCCAGCTCCGTTACAGACAGACCTCCAAATTTTAGGACAGCAGCA	4620
Db	4561	AGAGAAAAGTAAAGTCCAGCTCCGTTACAGACAGACCTCCAAATTTTAGGACAGCAGCA	4620
Qy	4621	TATGGATGGGATATTAAGGGGCTGAGACCTGAGAGCTGTCAAGATTTTCTCAACCCA	4680
Db	4621	TATGGATGGGATATTAAGGGGCTGAGACCTGAGAGCTGTCAAGATTTTCTCAACCCA	4680
Qy	4681	GCTAAGAGGAGTTTCGGGTGGGGCTCTTCAACCCACACAGACTCTCCCACTTAGAA	4740
Db	4681	GCTAAGAGGAGTTTCGGGTGGGGCTCTTCAACCCACACAGACTCTCCCACTTAGAA	4740
Qy	4741	GGAAATCTGCTTTCTTGGAAAGTGGGGTTCAGGCCGCTAGAGATCTGACAGGTGGCCTT	4800
Db	4741	GGAAATCTGCTTTCTTGGAAAGTGGGGTTCAGGCCGCTAGAGATCTGACAGGTGGCCTT	4800
Qy	4801	GGAAATCTGCTTTCTTGGAAAGTGGGGTTCAGGCCGCTAGAGATCTGACAGGTGGCCTT	4860
Db	4801	GGAAATCTGCTTTCTTGGAAAGTGGGGTTCAGGCCGCTAGAGATCTGACAGGTGGCCTT	4860

```

QY 4801 CCACGAGCTGGGAAGTTCTCAGTGGCAGGAGTTTCCACAAGAAAACATGAGGCCCT 4860
DB 4509 CCACGAGCTGGGAAGTTCTCAGTGGCAGGAGTTTCCACAAGAAAACATGAGGCCCT 4568
QY 4861 TCCCTTAAGCTGTCTTCTCCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4920
DB 4569 TCCCTTAAGCTGTCTTCTCCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4628
QY 4921 TGGGCTCTTGTCTTCAAGCAAGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4980
DB 4629 TGGGCTCTTGTCTTCAAGCAAGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4688
QY 4981 CTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5040
DB 4689 CTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4748
QY 5041 TTTTCTCCCTTCTCAATTGTAATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5100
DB 4749 TTTTCTCCCTTCTCAATTGTAATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4808
QY 5101 CTTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5160
DB 4809 CTTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4868
QY 5161 CTGCTGCAAGTGTGAGATCAACCTGGGGTTCACACCTTATGTAACATCTTCCA 5220
DB 4869 CTGCTGCAAGTGTGAGATCAACCTGGGGTTCACACCTTATGTAACATCTTCCA 4928
QY 5221 GTGAGCCACAGCTTCAAGTGTGCTGCTCTTCACTTCTCTCAACCCCTGGCTGTC 5280
DB 4929 GTGAGCCACAGCTTCAAGTGTGCTGCTCTTCACTTCTCTCAACCCCTGGCTGTC 4988
QY 5281 CTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5340
DB 4989 CTGCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5048
QY 5341 GCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5400
DB 5049 GCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5108
QY 5401 ATTCTTCAAGATTTCTGAAAGTTAAACAGTGAAGATTTTCCCTGTAGACAGCAGA 5460
DB 5109 ATTCTTCAAGATTTCTGAAAGTTAAACAGTGAAGATTTTCCCTGTAGACAGCAGA 5168
QY 5461 TCACGATTTCTCCGGAAGTCAAGGCTTCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5520
DB 5169 TCACGATTTCTCCGGAAGTCAAGGCTTCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5228
QY 5521 CTCTTGAACAACCTCAAGGCAACCTTACCCCAATAGACCTCTGACAGAGAAGAGCACT 5580
DB 5229 CTCTTGAACAACCTCAAGGCAACCTTACCCCAATAGACCTCTGACAGAGAAGAGCACT 5288
QY 5581 TTACATGAGTCTCTGCTGAGAGCCTAGGCTACGCTGTAAGAAAGAGCAGGAATGTGT 5640
DB 5289 TTACATGAGTCTCTGCTGAGAGCCTAGGCTACGCTGTAAGAAAGAGCAGGAATGTGT 5348
QY 5641 GGTGTGAGAAAGTCAAGACTTCAACATAGAAAGCTTACGCCACACCAAGATGACAGACAGA 5700
DB 5349 GGTGTGAGAAAGTCAAGACTTCAACATAGAAAGCTTACGCCACACCAAGATGACAGACAGA 5408
QY 5701 TCCCTCTATCTCCCAATAGAGTTGAGTGCAC 5735
DB 5409 TCCCTCTATCTCCCAATAGAGTTGAGTGCAC 5443

```

RESULT 3

```

US-10-332-966-2
; Sequence 2, Application US/10332966
; Publication No. US20030188324A1
; GENERAL INFORMATION:
; APPLICANT: HASEGAWA, Koji
; APPLICANT: KAWASE, Yosuke

```

```

; APPLICANT: SUZUKI, Hiroshi
; TITLE OF INVENTION: p300 TRANSGENIC ANIMAL
; FILE REFERENCE: 382.1040
; CURRENT APPLICATION NUMBER: US/10/332,966
; CURRENT FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: PCT/JP01/06086
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: JP2000-215143
; PRIOR FILING DATE: 2000-07-14
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 5443
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(5443)
US-10-332-966-2

```

```

Query Match      88.0%; Score 5045; DB 16; Length 5443;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 5383; Conservative 0; Mismatches 60; Indels 292; Gaps 1;

```

```

QY 1 GGATCCGTGCAAGTGCACACAGGGTCTCCACCCACAGGTCCTAGTCTCAATTTCACT 60
DB 1 GGATCCGTGCAAGTGCACACAGGGTCTCCACCCACAGGTCCTAGTCTCAATTTCACT 60
QY 61 TTCCATGCTTGTCTCACAATGTGCTCTCCCAAGCTTAATTTGACTTTGTTTAT 120
DB 61 TTCCATGCTTGTCTCACAATGTGCTCTCCCAAGCTTAATTTGACTTTGTTTAT 120
QY 121 TTCAAAAGGCTGATGAGAGTATGTTGTGCTACCACTCTTAAGGGGCCGCTGA 180
DB 121 TTCAAAAGGCTGATGAGAGTATGTTGTGCTACCACTCTTAAGGGGCCGCTGA 180
QY 181 AGCCCTCAGACCTGAGGCTTTGCAACAGCCCTTAAAGTGAAGCAGATTAAGCAATT 240
DB 181 AGCCCTCAGACCTGAGGCTTTGCAACAGCCCTTAAAGTGAAGCAGATTAAGCAATT 240
QY 241 TCCCTTAAGCAAAATCTGCTCTAGACTCTTCTCTCTCTGACTCTGCTGCTGCTGCT 300
DB 241 TCCCTTAAGCAAAATCTGCTCTAGACTCTTCTCTCTCTGACTCTGCTGCTGCTGCT 300
QY 301 AGGCTGGGAGGCTGGGCTTGAAGAAAGAGTGGGAAAGTGGCAAAACCGATCCCTAG 360
DB 301 AGGCTGGGAGGCTGGGCTTGAAGAAAGAGTGGGAAAGTGGCAAAACCGATCCCTAG 360
QY 361 GGCCCTGTAGAGTGTGAGGCTTCCCTGTACAGCACTGCTCATAGATCTCTCCAGCC 420
DB 361 GGCCCTGTAGAGTGTGAGGCTTCCCTGTACAGCACTGCTCATAGATCTCTCCAGCC 420
QY 421 AAACATAGCAAGATGATACCTCTTGTGTGACTTCCAGAGCCCAAGTACCTGTAGGTT 480
DB 421 AAACATAGCAAGATGATACCTCTTGTGTGACTTCCAGAGCCCAAGTACCTGTAGGTT 480
QY 481 GAAACAGATTTTGAAGAGCCTTGAACCTCACTGAACCTCTGAAGTCTATCCACCAAGA 540
DB 481 GAAACAGATTTTGAAGAGCCTTGAACCTCACTGAACCTCTGAAGTCTATCCACCAAGA 540
QY 541 AGCAGCTAGTGCACCTGCTAGTATGATCTTACGCTGATTAATATGACAGAGTGGCCAC 600
DB 541 AGCAGCTAGTGCACCTGCTAGTATGATCTTACGCTGATTAATATGACAGAGTGGCCAC 600
QY 601 AGAAGTCTGGGCTGTAGAACTGACCACTGACTTTTCACTGGGCAAGGTATGACCCCC 660
DB 601 AGAAGTCTGGGCTGTAGAACTGACCACTGACTTTTCACTGGGCAAGGTATGACCCCC 660
QY 661 TCAGCAGATGTAGTATGTCCTCTTGAATCCATCCAGGCAAGTCTCTTAAGAGCAGT 720
DB 661 TCAGCAGATGTAGTATGTCCTCTTGAATCCATCCAGGCAAGTCTCTTAAGAGCAGT 720
QY 721 GGATGAGAGATGTAGTATGATGGAATTCCAAAACAGATATCCAGATGTCCTTGCCTCC 780

```


Db 721 GGATGAGAGATGATGATGCGATTCACAAACACACTATCCACATGTCCTTGCCCC 780
Qy 781 TTCACCTTAGCAGAGAGACAGTAACCTTAGCCTATCTTCTCTCCATCCTCCAG 840
Db 781 TTCACCTTAGCAGAGAGACAGTAACCTTAGCCTATCTTCTCTCCATCCTCCAG 840
Qy 841 GACACACCCCTGCTGTCAGATATTCATTTCTTCCAGTCCTCTGTGACTTCAT 900
Db 841 GACACACCCCTGCTGTCAGATATTCATTTCTTCCAGTCCTCTGTGACTTCAT 900
Qy 901 TTGGAAAGCTTTTGAACCTCTGACAGCTGCTGGAAGATAGATTTGGCCTTAGTGGCAA 960
Db 901 TTGGAAAGCTTTTGAACCTCTGACAGCTGCTGGAAGATAGATTTGGCCTTAGTGGCAA 960
Qy 961 GCCATCTCAAGAGAAACAGACACAGGGGAGCCAGATTTTGGAGAGATCAGAACTAAA 1020
Db 961 GCCATCTCAAGAGAAACAGACACAGGGGAGCCAGATTTTGGAGAGATCAGAACTAAA 1020
Qy 1021 TCACCTGGCGGGCTTGCGGGGTAGAAAAAAGATGAGTGCCTCCGCTCAGCTTAAGCCAGC 1080
Db 1021 TCACCTGGCGGGCTTGCGGGGTAGAAAAAAGATGAGTGCCTCCGCTCAGCTTAAGCCAGC 1080
Qy 1081 TAGTCCCCGAGATCTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGTTC 1140
Db 1081 TAGTCCCCGAGATCTCTGCCACAGCTGGGCTGCTCGGGGTAGCTTTAGGAATGTGGTTC 1140
Qy 1141 TGAAGAACAATGGGATTTGGAAGACATCTTTTGAAGTCTCCCTCAACCCCACTACAGAC 1200
Db 1141 TGAAGAACAATGGGATTTGGAAGACATCTTTTGAAGTCTCCCTCAACCCCACTACAGAC 1200
Qy 1201 ACACCTGTGTGGCCAGACTCTGTTTCAACAGCCCTCTGTGTCTACCACTGAGCTAG 1260
Db 1201 ACACCTGTGTGGCCAGACTCTGTTTCAACAGCCCTCTGTGTCTACCACTGAGCTAG 1260
Qy 1261 GCAACCAAGAGATGGGCTGTGCTGAGATGGAAGCTGGTTACCAATGCAAAAACAG 1320
Db 1261 GCAACCAAGAGATGGGCTGTGCTGAGATGGAAGCTGGTTACCAATGCAAAAACAG 1320
Qy 1321 CAGGGAGGGAGAAACAGAGAACGAAATAGAGAGAGAAAGGAGCCAGTCAATCAGA 1380
Db 1321 CAGGGAGGGAGAAACAGAGAACGAAATAGAGAGAGAAAGGAGCCAGTCAATCAGA 1380
Qy 1381 TGCACTCAGAAAGATGGAAAGCCACACACAGCTTGAAGCAGAGAAACAGAAAAGGAG 1440
Db 1381 TGCACTCAGAAAGATGGAAAGCCACACACAGCTTGAAGCAGAGAAACAGAAAAGGAG 1440
Qy 1441 AGATTTCTGGGATTAAGAGGCGACAGAAAGAGCCAGGCCCCCAAGTCTCTCTTT 1500
Db 1441 AGATTTCTGGGATTAAGAGGCGCGACAGAAAGAGCCAGGCCCCCAAGTCTCTCTTT 1500
Qy 1501 ATACCTCATCCCGTCTCCCAATTAAAGCCCACTCTTCTCTAGATCAGACCTGAGCTGC 1560
Db 1501 ATACCTCATCCCGTCTCCCAATTAAAGCCCACTCTTCTCTAGATCAGACCTGAGCTGC 1560
Qy 1561 AGCGAAGAGACCCGTAAGAGAGATCACTAGATGAAGAGATGTGTGAGAAATGCCAGG 1620
Db 1561 AGCGAAGAGACCCGTAAGAGAGATCACTAGATGAAGAGATGTGTGAGAAATGCCAGG 1620
Qy 1621 GCAACCTTAAGAGAGAGGCTTAAGAGAGAGATTAAGTGTCTTCAAAAGTGGCCAGG 1680
Db 1621 GCAACCTTAAGAGAGAGGCTTAAGAGAGAGATTAAGTGTCTTCAAAAGTGGCCAGG 1680
Qy 1681 CTGTGACACAGAGGCTGAGAGCTGTGTGAGAGCTCAAGATTAAGATGATGCTCAGA 1740
Db 1681 CTGTGACACAGAGGCTGAGAGCTGTGTGAGAGCTCAAGATTAAGATGATGCTCAGA 1740
Qy 1741 ATGGAGCGGGGGGGGATTTCTGGGGGGGGGAGAGAGAGTGAAGAGAGCTTGAAACAG 1800
Db 1741 ATGGAGCGGGGGGGGATTTCTGGGGGGGGGAGAGAGAGTGAAGAGAGCTTGAAACAG 1800
Qy 1801 AGAATTCGAGAGCCCTGGAACGATACCATAAAGGAGAAACCAAGGCTACTTTAATG 1860
Db 1801 AGAATTCGAGAGCCCTGGAACGATACCATAAAGGAGAAACCAAGGCTACTTTAATG 1860

Db 1801 AGAATTCGAGAGCCCTGGAACGATACCATAAAGGAGAAACCAAGGCTACTTTAATG 1860
Qy 1861 TAAATCATGAAAGACAGGAGAGAGGGAGCTGAGAGATGAGAGAGAGCCCGGGCAGAG 1920
Db 1861 TAAATCATGAAAGACAGGAGAGAGGGAGCTGAGAGATGAGAGAGAGCCCGGGCAGAG 1920
Qy 1921 CATGAGAGAGAGACAGGAGGTTGAGGCTTCCTGAAATCAGCTGCTGAAGGAGAG 1980
Db 1921 CATGAGAGAGAGACAGGAGGTTGAGGCTTCCTGAAATCAGCTGCTGAAGGAGAG 1980
Qy 1981 CCTGGTATGAGACCCAGAAACAGAGAGGCTAGGGTTAATGTGAGACAGGGAAACAGAG 2040
Db 1981 CCTGGTATGAGACCCAGAAACAGAGAGGCTAGGGTTAATGTGAGACAGGGAAACAGAG 2040
Qy 2041 GTAGACACAGAAACAGACAGAGACGGGGAGCCAGGTAAACAAAGAAATGCTCTTCAAC 2100
Db 2041 GTAGACACAGAAACAGACAGAGACGGGGAGCCAGGTAAACAAAGAAATGCTCTTCAAC 2100
Qy 2101 CTGTGGCCAGAGGCTCATCTGTGTCAACATCTTGAATGTTCACTCAGACTGACGGGC 2160
Db 2101 CTGTGGCCAGAGGCTCATCTGTGTCAACATCTTGAATGTTCACTCAGACTGACGGGC 2160
Qy 2161 TGGCTTTGGAGGCGCTGGAAGAAGTATGTGAGAGCAGGGAGACAGAGGGGCTTAGGA 2220
Db 2161 TGGCTTTGGAGGCGCTGGAAGAAGTATGTGAGAGCAGGGAGACAGAGGGGCTTAGGA 2220
Qy 2221 AAGAGAGAGGGGCAACAGGCAACAGAGAGGAGAGGAGCCAGACTGAGTTAACTC 2280
Db 2221 AAGAGAGAGGGGCAACAGGCAACAGAGAGGAGAGGAGCCAGACTGAGTTAACTC 2280
Qy 2281 CTTCCTTTTGTGATCTTCCATAGAGAGCAGTGGAACTCTGTGACCAACATCCCATATA 2340
Db 2281 CTTCCTTTTGTGATCTTCCATAGAGAGCAGTGGAACTCTGTGACCAACATCCCATATA 2340
Qy 2341 GCCCCCACTACCATACCAAGTTTGGCTGAGTGGCATTTAGTTCCTCTGAGAGACAGAG 2400
Db 2341 GCCCCCACTACCATACCAAGTTTGGCTGAGTGGCATTTAGTTCCTCTGAGAGACAGAG 2400
Qy 2401 CCTGGCTTTTGTCTCTTGAACCTGACCTGACCCAAAGTTTCTCAGTACTTTGAAT 2460
Db 2401 CCTGGCTTTTGTCTCTTGAACCTGACCCAAAGTTTCTCAGTACTTTGAAT 2460
Qy 2461 GCCCTCAAGCTTGAAGAACAGGAGAGTGAATTAAGCCATGGGCTAAACCTTGAAGCT 2520
Db 2461 GCCCTCAAGCTTGAAGAACAGGAGAGTGAATTAAGCCATGGGCTAAACCTTGAAGCT 2520
Qy 2521 TGCAACAGAGGCTCAAGTGAACCTCAGGGAGACAGCTGACAGAGGTGGCTTTATCC 2580
Db 2521 TGCAACAGAGGCTCAAGTGAACCTCAGGGAGACAGCTGACAGAGGTGGCTTTATCC 2580
Qy 2581 CCAAGAGCAACCATTTGGCATAGTGTGCTGCAAAATGGGAATGCAAGTTGAATCAGGTC 2640
Db 2581 CCAAGAGCAACCATTTGGCATAGTGTGCTGCAAAATGGGAATGCAAGTTGAATCAGGTC 2640
Qy 2641 CCTTCAAGAAATCTGCAATGCAAGACCTTAAGACCCCTGAGAGAGGGGTATGCTCTGCC 2700
Db 2641 CCTTCAAGAAATCTGCAATGCAAGACCTTAAGACCCCTGAGAGAGGGGTATGCTCTGCC 2700
Qy 2701 CCAACCACTAAGGGGAGTGAACCTATCTAAGGGGCTGGCGACTTTGGGAGACACAC 2760
Db 2701 CCAACCACTAAGGGGAGTGAACCTATCTAAGGGGCTGGCGACTTTGGGAGACACAC 2760
Qy 2761 ATTACTGAGAGTCTGAGCCAGAAAAAAGTGAACCGCCCTGTGTCTTGCCCACTCAAC 2820
Db 2761 ATTACTGAGAGTCTGAGCCAGAAAAAAGTGAACCGCCCTGTGTCTTGCCCACTCAAC 2820
Qy 2821 TCTAGAGCTATATTGAGAGGTGAACATTAAGTATAGGGTGGAGCTGTGAGAGAGAGTGT 2880
Db 2821 TCTAGAGCTATATTGAGAGGTGAACATTAAGTATAGGGTGGAGCTGTGAGAGAGAGTGT 2880
Qy 2881 TCCGTGGGTGAGAGGTGTAGGGGAAAGCCAGACAGAGGAGTCTGGCTTTGTCTCTGAA 2940
Db 2881 TCCGTGGGTGAGAGGTGTAGGGGAAAGCCAGACAGAGGAGTCTGGCTTTGTCTCTGAA 2940

2941 CCAATGTCTACTTACTTAAACAGGATGACCTGCTAAAGACCCCAACATCTACGACCTC 3000
2941 CCAATGTCTACTTACTTAAACAGGATGACCTGCTAAAGACCCCAACATCTACGACCTC 3000
3001 TGAAGAGACGACGACCTGAGAGACAGGGGTTGTCTCTGAGCCTTGGGTCTTGATGTG 3060
3001 TGAAGAGACGACGACCTGAGAGACAGGGGTTGTCTCTGAGCCTTGGGTCTTGATGTG 3060
3061 CCAAGAGAGAGGAGCATGAGTGTGATTAAGGCCCAAGAGCGTTAGAGAGGCACTTG 3120
3061 CCAAGAGAGAGGAGCATGAGTGTGATTAAGGCCCAAGAGCGTTAGAGAGGCACTTG 3120
3121 GGAAGGGGTCACTGCTGACAGAGCCCTATCAGGATCTGAGGCTGGGGCCAACTGTG 3180
3121 GGAAGGGGTCACTGCTGACAGAGCCCTATCAGGATCTGAGGCTGGGGCCAACTGTG 3180
3181 TAAATCTGAGGCTGACAGGACATTCAGAGCAGCATCTGATCTCTGAGCCTGGGGA 3240
3181 TAAATCTGAGGCTGACAGGACATTCAGAGCAGCATCTGATCTCTGAGCCTGGGGA 3240
3241 GGGGGAAGGAGCAACCCCACTTATACCTTTCTCTCTCAGGCCAGAGATTACACT 3300
3241 GGGGGAAGGAGCAACCCCACTTATACCTTTCTCTCTCAGGCCAGAGATTACACT 3300
3301 CTGGGCTTCCCTCCCTCCACCTCCATCAGAGTGAAGGAGGTGCAAGAGGATTAATA 3360
3301 CTGGGCTTCCCTCCCTCCACCTCCATCAGAGTGAAGGAGGTGCAAGAGGATTAATA 3360
3361 CCTACATGTCACAAATCATGTGTGACAGATATATGATCAGTATGTGTAGAGCAAGAA 3420
3361 CCTACATGTCACAAATCATGTGTGACAGATATATGATCAGTATGTGTAGAGCAAGAA 3420
3421 GGAATCTGAGGCTTAACTGGTTATGTGTAAAGTCTGTGTGATGTGTGTCTG 3480
3421 GGAATCTGAGGCTTAACTGGTTATGTGTAAAGTCTGTGTGATGTGTGTCTG 3480
3481 ACTGAAAGGGGATGAGCTGTGACAGTGTGAGTCTGTGAGGATTAACAGATGCA 3540
3481 ACTGAAAGGGGATGAGCTGTGACAGTGTGAGTCTGTGAGGATTAACAGATGCA 3540
3541 GGTGTGTGTAAATTTGCAAGGCAAGGAGGATGAGTCTCTCAATGTTTAAAGAT 3600
3541 GGTGTGTGTAAATTTGCAAGGCAAGGAGGATGAGTCTCTCAATGTTTAAAGAT 3600
3601 TGGATGATGAGCTGATCTCAAGGACATGAGAAATAGAAATGACACTATATGTGTCT 3660
3601 TGGATGATGAGCTGATCTCAAGGACATGAGAAATAGAAATGACACTATATGTGTCT 3660
3661 CTAAGCTAAGTAAAGGATCTTTTGAAGACACTGTCTTAAGATGTGGGCAACAGAGC 3720
3661 CTAAGCTAAGTAAAGGATCTTTTGAAGACACTGTCTTAAGATGTGGGCAACAGAGC 3720
3721 TACAGACAGTATCTGACAGATTAAGAGAGAGAGGAGGAGTGTGAATTTCTTACTA 3780
3721 TACAGACAGTATCTGACAGATTAAGAGAGAGAGGAGGAGTGTGAATTTCTTACTA 3780
3781 TCAAGGAGAACTGATCTGACACTGCAAAAGTGAATGCTCTCTTGAACATGACTT 3840
3781 TCAAGGAGAACTGATCTGACACTGCAAAAGTGAATGCTCTCTTGAACATGACTT 3840
3841 TGTCTCTGGGAGCCAGCACTGTGAACTTCAAGTCTGAGAGTGAAGGCTCCCTCA 3900
3841 TGTCTCTGGGAGCCAGCACTGTGAACTTCAAGTCTGAGAGTGAAGGCTCCCTCA 3900
3901 GCTTGAAGTATGACATAGCAGGTTGAAAGGAGGAGAGGAGGAGGAGGAGGAGGAGG 3960
3901 GCTTGAAGTATGACATAGCAGGTTGAAAGGAGGAGAGGAGGAGGAGGAGGAGGAGG 3960
3961 TTTGTGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4020
3961 TTTGTGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4020

4021 TTGTTCAACTACCTCTCAGATTAATAATACTGAGTAAAGGCTGGGTAGGGAGGTG 4080
4021 TTGTTCAACTACCTCTCAGATTAATAATACTGAGTAAAGGCTGGGTAGGGAGGTG 4080
4081 GTGTGAGAGGCTCTGTCTCTCTCTGATGCTTGAAGGCTTTGGGAGAGGAATGT 4140
4081 GTGTGAGAGGCTCTGTCTCTCTCTGATGCTTGAAGGCTTTGGGAGAGGAATGT 4140
4141 GCCCAAGAGCTTAATAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4200
4141 GCCCAAGAGCTTAATAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4200
4201 GCAAACTTGGGGCCCTGATGATGATTAAGAACTGCGCAATGATACCTTCTTC 4260
4201 GCAAACTTGGGGCCCTGATGATGATTAAGAACTGCGCAATGATACCTTCTTC 4260
4261 GCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4320
4261 GCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4320
4321 AAGTGAAGTTAACCACTCCCTATCAGTATAGAGAAAGTGAAGTGAAGTTTACACTC 4380
4321 AAGTGAAGTTAACCACTCCCTATCAGTATAGAGAAAGTGAAGTGAAGTTTACACTC 4380
4381 CCTATCAGTATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4440
4381 CCTATCAGTATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4440
4441 AAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4500
4441 AAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4500
4501 CCACTCCCTATCAGTATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4560
4501 CCACTCCCTATCAGTATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4560
4561 AGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4620
4561 AGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4620
4621 TATGGATGAGTATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4680
4621 TATGGATGAGTATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4680
4681 GGTAAAGAGGAGTGTGGGTGGGGCTCTTCAACCAACAGAGCTTCTCCCACTAGAA 4740
4681 GGTAAAGAGGAGTGTGGGTGGGGCTCTTCAACCAACAGAGCTTCTCCCACTAGAA 4740
4741 GGAAGCTGCTTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4800
4741 GGAAGCTGCTTCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4800
4801 CCACAGGCTGGAGGAGTCTCAGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4860
4801 CCACAGGCTGGAGGAGTCTCAGTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4860
4861 TCCCTTAAGGCTCTCTTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4920
4861 TCCCTTAAGGCTCTCTTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4920
4921 TGGGCT 4980
4921 TGGGCT 4980
4981 CTCGCT 5040
4981 CTCGCT 5040
5041 TTTTCT 5100
5041 TTTTCT 5100
5101 CTTCCT 5160

[illegible]

```

RESULT 4
US-10-944-375-15
; Sequence 15, Application US/10944375
; Publication No. US2005006381A1
; GENERAL INFORMATION:
; APPLICANT: Cincinnati Children's Hospital Medical Center
; APPLICANT: University of Cincinnati
; APPLICANT: Molkenstin, Jeffery D
; APPLICANT: Kraniak, Evangelia G
; TITLE OF INVENTION: REGULATION OF CARDIAC CONTRACTILITY AND HEART FAILURE PROPENSITY
; FILE REFERENCE: 9761M#L
; CURRENT APPLICATION NUMBER: US/10/944,375
; CURRENT FILING DATE: 2004-09-17
; PRIOR APPLICATION NUMBER: US 60/503,853
; PRIOR FILING DATE: 2003-09-19
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 15
; LENGTH: 5443
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-944-375-15

```

Query Match	88.0%;	Score 5045;	DB 21;	Length 5443;
Best Local Similarity	93.9%;	Pred. No. 0;		
Matches 5383;	Conservative	0;	Mismatches 60;	Indels 292;
				Gaps 1;

OY 1 GGATCTGCAGGTACACACAGGGGTCCTCCACCACAGGTGCGCCCTAGTCTCAATTTCGT 60

D	b	1	GGATCTCGCAAGGTCAACACAAAGGCTCTCCACCCACAGAGTGGCCCTAATCTCAATTACGT	60
Q	y	61	TTCCATGCTGTGTTTCTCACATGCTGGCTCCCCAGAGCTAATTGGACTTTGTTTTAT	120
D	b	61	TTCCATGCTGTGTTTCTCACATGCTGGCTCCCCAGAGCTAATTGGACTTTGTTTTAT	120
Q	y	121	TTCAAAAGGGCTTGAAATGAGAGTATGATCTTTGGTCTATCCAGCTCTAAGGGTGGCGTGA	180
D	b	121	TTCAAAAGGGCTTGAAATGAGAGTATGATCTTTGGTCTATCCAGCTCTAAGGGTGGCGTGA	180
Q	y	181	AGCCCTCAGACCTGGAGCCCTTTGCCAACAGCCCTTTAGGTGAAGACAAATPAAGCAATTT	240
D	b	181	AGCCCTCAGACCTGGAGCCCTTTGCCAACAGCCCTTTAGGTGAAGACAAATPAAGCAATTT	240
Q	y	241	TCCTTAAAGCCAAATCTGCTCTAGACTCTTCTTCTGACCTCGTCCCTGGCTCT	300
D	b	241	TCCTTAAAGCCAAATCTGCTCTAGACTCTTCTTCTGACCTCGTCCCTGGCTCT	300
Q	y	301	AGGTTGGGGAGGTGGGGCTTGGAAGAAAGTGGGGAAATGGGAAAAGCCGATCCCTAG	360
D	b	301	AGGTTGGGGAGGTGGGGCTTGGAAGAAAGTGGGGAAATGGGAAAAGCCGATCCCTAG	360
Q	y	361	GGCCCTGTGAAGTTGGAGCCCTTCCGTGACAGCACTGGCTATGATTCCTCTCCAGCC	420
D	b	361	GGCCCTGTGAAGTTGGAGCCCTTCCGTGACAGCACTGGCTATGATTCCTCTCCAGCC	420
Q	y	421	AAACATAGCAAGAAGTATACCTCTCTTGTGACTCTTCCAGGCCAGTACTGTCAAGTT	480
D	b	421	AAACATAGCAAGAAGTATACCTCTCTTGTGACTCTTCCAGGCCAGTACTGTCAAGTT	480
Q	y	481	GAAACAGATTTTAGAGAAAGCTCTGAACTCACTGAACTCTGAAAGTCTATCCACCAAGCA	540
D	b	481	GAAACAGATTTTAGAGAAAGCTCTGAACTCACTGAACTCTGAAAGTCTATCCACCAAGCA	540
Q	y	541	AGCACCTAAGGTGCCACCTGATAGTATGATCCTAAGCTGATTAATGAGAGAGCTGGGCAC	600
D	b	541	AGCACCTAAGGTGCCACCTGATAGTATGATCCTAAGCTGATTAATGAGAGAGCTGGGCAC	600
Q	y	601	AGAACTCCTGGGGTGTAGAACTGACCAAGTGAATCTTTCACTGGCAAAAGTATGACCCC	660
D	b	601	AGAACTCCTGGGGTGTAGAACTGACCAAGTGAATCTTTCACTGGCAAAAGTATGACCCC	660
Q	y	661	TCAGCAGATGTATATGTCTCCCTTAAATCCCATCCCAAGGCTCTTAAGAGACATG	720
D	b	661	TCAGCAGATGTATATGTCTCCCTTAAATCCCATCCCAAGGCTCTTAAGAGACATG	720
Q	y	721	GGATGAGAGATGTATGTATGTGGCATTCGAAACAGGTATCCACAGTGTCCCTTGCCCC	780
D	b	721	GGATGAGAGATGTATGTATGTGGCATTCGAAACAGGTATCCACAGTGTCCCTTGCCCC	780
Q	y	781	TTCCACTTAGCCAGAGAGACAGTAACTTAGCCTATCTTCTCTCCCATCTCTCCAG	840
D	b	781	TTCCACTTAGCCAGAGAGACAGTAACTTAGCCTATCTTCTCTCCCATCTCTCCAG	840
Q	y	841	GACACACCCCTGGCTGACAGTATTCATTTCTCTTCAAGTCCCTCTGTGACTTCCAT	900
D	b	841	GACACACCCCTGGCTGACAGTATTCATTTCTCTTCAAGTCCCTCTGTGACTTCCAT	900
Q	y	901	TTTGCAGAGCTTTTAACTCTGACACTGTGGAAATATGAGTTTGACCTTAAGTGTGGCAA	960
D	b	901	TTTGCAGAGCTTTTAACTCTGACACTGTGGAAATATGAGTTTGACCTTAAGTGTGGCAA	960
Q	y	961	GCATCTCTCAAGAGAAAGACAGCAACAGGGGGACAGATTTTGGAAAGATTCAGAACTTAA	1020
D	b	961	GCATCTCTCAAGAGAAAGACAGCAACAGGGGGACAGATTTTGGAAAGATTCAGAACTTAA	1020
Q	y	1021	TCACTGGCGGGCTGGGGGTGAAAAAAGATGATGATCGCTCCAGCTTAAGCCAAAGC	1080
D	b	1021	TCACTGGCGGGCTGGGGGTGAAAAAAGATGATGATCGCTCCAGCTTAAGCCAAAGC	1080
Q	y	1081	TATGTCCTCCGAGATATCTGTGCAACAGCTGGGCTGCTCGGGGTAGACTTTTAGAAATGTGGGT	1140
D	b	1081	TATGTCCTCCGAGATATCTGTGCAACAGCTGGGCTGCTCGGGGTAGACTTTTAGAAATGTGGGT	1140

QY 1141 TGAAGCAATGGGATTGGAAGCATCTCTTTGAGTCTCCCTCAACCCCACTAGAGAC 1200
DB 1141 TGAAGCAATGGGATTGGAAGCATCTCTTTGAGTCTCCCTCAACCCCACTAGAGAC 1200
QY 1201 ACACTGTGTGTGGCAGACTCCTGTTCAACAGCCCTGTGTGTTGACCACTGAGCTAG 1260
DB 1201 ACACTGTGTGTGGCAGACTCCTGTTCAACAGCCCTGTGTGTTGACCACTGAGCTAG 1260
QY 1261 GCACCAAGCATGGGCTGTGTCTGAGATGAAAGTTGGTTTCAATTCGAAAAACAG 1320
DB 1261 GCACCAAGCATGGGCTGTGTCTGAGATGAAAGTTGGTTTCAATTCGAAAAACAG 1320
QY 1321 CAGGGAGGGGAGAACAGAAAGAAATTAAGGAAGAAAGGAAAGGCAAGTCAATCAG 1380
DB 1321 CAGGGAGGGGAGAACAGAAAGAAATTAAGGAAGAAAGGAAAGGCAAGTCAATCAG 1380
QY 1381 TGCAGTCAGAAAGATGGGAAGCCAAACACAGCTTGAGCAGAGAAACAGAAAAAGGAG 1440
DB 1381 TGCAGTCAGAAAGATGGGAAGCCAAACACAGCTTGAGCAGAGAAACAGAAAAAGGAG 1440
QY 1441 AGATTCTGGCATTAAGAGGCCACAGAAAGAGGCCAGGCCCAAGTCTCTCTTT 1500
DB 1441 AGATTCTGGCATTAAGAGGCCACAGAAAGAGGCCAGGCCCAAGTCTCTCTTT 1500
QY 1501 ATACCCCATCCCGTCTCCCAATTAAAGCCCATCTTCTTCTAGATCAGACCTGAGCTGC 1560
DB 1501 ATACCCCATCCCGTCTCCCAATTAAAGCCCATCTTCTTCTAGATCAGACCTGAGCTGC 1560
QY 1561 AGCGAAGAGACCCGTAGAGAGATCACTGATGAAAGAGATGTGTGAGAAAGTCCAG 1620
DB 1561 AGCGAAGAGACCCGTAGAGAGATCACTGATGAAAGAGATGTGTGAGAAAGTCCAG 1620
QY 1621 GCACCTTAAGAGCTCAAGCCTTAAAGAGCAAGATTAAGGTCTTCAAGGTGGCCAG 1680
DB 1621 GCACCTTAAGAGCTCAAGCCTTAAAGAGCAAGATTAAGGTCTTCAAGGTGGCCAG 1680
QY 1681 CTGTGACACAGAGGCTGAGAGCTGTGTGTAAGGCTCAAGTAAGATGATGCTCAG 1740
DB 1681 CTGTGACACAGAGGCTGAGAGCTGTGTGTAAGGCTCAAGTAAGATGATGCTCAG 1740
QY 1741 ATGGGCGGGGGGGGATTTCTGGGGGGGGGAGAGAAAGGTGAGAAAGGCTGGAACAG 1800
DB 1741 ATGGGCGGGGGGGGATTTCTGGGGGGGGGAGAGAAAGGTGAGAAAGGCTGGAACAG 1800
QY 1801 AGAATCTGGAAGCGCTGGAAGCAATACATTAAGGGAAGACCAGGCTACCTTTAGATG 1860
DB 1801 AGAATCTGGAAGCGCTGGAAGCAATACATTAAGGGAAGACCAGGCTACCTTTAGATG 1860
QY 1861 TAAATCATGAAGACAGGGAAGGGAAGCTGAGAGAGTGAAGAACCCCGGGGCAAGA 1920
DB 1861 TAAATCATGAAGACAGGGAAGGGAAGCTGAGAGAGTGAAGAACCCCGGGGCAAGA 1920
QY 1921 CATGAAGCAAGGACCAAGGCTGAGCGCTCCGTGAATCAGCCTGTGGAAGCAGAG 1980
DB 1921 CATGAAGCAAGGACCAAGGCTGAGCGCTCCGTGAATCAGCCTGTGGAAGCAGAG 1980
QY 1981 CCTGTGATGAGACCAAGCAGCAGAGGCTAGGGTTAATGTGAGACAGGGAACAGAG 2040
DB 1981 CCTGTGATGAGACCAAGCAGCAGAGGCTAGGGTTAATGTGAGACAGGGAACAGAG 2040
QY 2041 GTTGAACAGGGAACAGACAGAGCGGGGAGCAAGGTAACAAGGAATGTTCTTCTCAC 2100
DB 2041 GTTGAACAGGGAACAGACAGAGCGGGGAGCAAGGTAACAAGGAATGTTCTTCTCAC 2100
QY 2101 CTGTGGCAGAGGCTCATCTGTCTCACTACTTGAATGTTTCATCAGACTGAGGGC 2160
DB 2101 CTGTGGCAGAGGCTCATCTGTCTCACTACTTGAATGTTTCATCAGACTGAGGGC 2160
QY 2161 TGGCTTGGAGGCACTGGAAGATGTGAGAGCCAGGGGAGACAAAGGGGCTTAGGA 2220
DB 2161 TGGCTTGGAGGCACTGGAAGATGTGAGAGCCAGGGGAGACAAAGGGGCTTAGGA 2220

QY 2221 AAGAAAGAGGGGCAACCAAGGCCACAGAGGGCAGAGGCCAGAACTGATTAAGTCTC 2280
DB 2221 AAGAAAGAGGGGCAACCAAGGCCACAGAGGGCAGAGGCCAGAACTGATTAAGTCTC 2280
QY 2281 CTTCTGTGTGATCTTCATAGAGAGGTGGAACTGTGACCAACATCCCATGA 2340
DB 2281 CTTCTGTGTGATCTTCATAGAGAGGTGGAACTGTGACCAACATCCCATGA 2340
QY 2341 GCGCCCACTAACCAATGAAAGTTGGGCTGAGTGCATCTAGTTCCCTGAGGAAGAG 2400
DB 2341 GCGCCCACTAACCAATGAAAGTTGGGCTGAGTGCATCTAGTTCCCTGAGGAAGAG 2400
QY 2401 CTTGGCTTTGTCTCTGAGACCTGACCCAGGCTGACCCATGTTCTCACTACTTGAAT 2460
DB 2401 CTTGGCTTTGTCTCTGAGACCTGAGCCTGAGGCTGAGGCTTAACTTACTTCAAT 2460
QY 2461 GCGCTCAAGAGCTTGAAGCAAGGAGTGAATTAAGGCAATGGCTTAACTTGAAGCT 2520
DB 2461 GCGCTCAAGAGCTTGAAGCAAGGAGTGAATTAAGGCAATGGCTTAACTTGAAGCT 2520
QY 2521 TGCACACAGAGGCTCAAGTGAACCTCAGGAGCAGAGCTGACAGAGGCTTATCC 2580
DB 2521 TGCACACAGAGGCTCAAGTGAACCTCAGGAGCAGAGCTGACAGAGGCTTATCC 2580
QY 2581 CCAAGAGCAACCATTTGATAGGTGCTGCAAAATGGGAATGCAAGTTGAATCAGTTC 2640
DB 2581 CCAAGAGCAACCATTTGATAGGTGCTGCAAAATGGGAATGCAAGTTGAATCAGTTC 2640
QY 2641 CTTTCAAGAAATCTGATGCAAGACCTTAAGACCCCTGAGAGAGGGTAATGCTCTGCC 2700
DB 2641 CTTTCAAGAAATCTGATGCAAGACCTTAAGACCCCTGAGAGAGGGTAATGCTCTGCC 2700
QY 2701 CCAACCACTAAGGGAATGAATCTTCTAGGGGCTGGCCACTTTGGGAGACACAC 2760
DB 2701 CCAACCACTAAGGGAATGAATCTTCTAGGGGCTGGCCACTTTGGGAGACACAC 2760
QY 2761 ATTAAGAGAGTGTAGGCGCCAGAAAACTGACCGGCTGTGTCTGACCACCTCCACAC 2820
DB 2761 ATTAAGAGAGTGTAGGCGCCAGAAAACTGACCGGCTGTGTCTGACCACCTCCACAC 2820
QY 2821 TCTAGAGCTAATTAAGAGGTGACATGATTAAGGTGGAGCTGTAGCAGGAGAGTGT 2880
DB 2821 TCTAGAGCTAATTAAGAGGTGACATGATTAAGGTGGAGCTGTAGCAGGAGAGTGT 2880
QY 2881 TCTGTGTGTGAGGCTGTAGGGGAAAGCCAGACAGGGAAGTGTGCTTGTCTTGA 2940
DB 2881 TCTGTGTGTGAGGCTGTAGGGGAAAGCCAGACAGGGAAGTGTGCTTGTCTTGA 2940
QY 2941 CACAATGTCTACTAGTTAATTAAGGATGACCTGCTAAGACCCCAATCTACGACTTC 3000
DB 2941 CACAATGTCTACTAGTTAATTAAGGATGACCTGCTAAGACCCCAATCTACGACTTC 3000
QY 3001 TGAAGAGCAGAGCCCTGAGAGACAGGGGTGTCTCTAGGCTTGGGTCTTGAATGTG 3060
DB 3001 TGAAGAGCAGAGCCCTGAGAGACAGGGGTGTCTCTAGGCTTGGGTCTTGAATGTG 3060
QY 3061 CCAAAAGAGGCGCATGATGTGATTAAGGCCCAAGAGGTTTGAAGAGGCACTTG 3120
DB 3061 CCAAAAGAGGCGCATGATGTGATTAAGGCCCAAGAGGTTTGAAGAGGCACTTG 3120
QY 3121 GGAAGGGGTCAGTCTGACAGGCCCTTATCATGGAATCTGAGGCTGGGGCCAACTGGTG 3180
DB 3121 GGAAGGGGTCAGTCTGACAGGCCCTTATCATGGAATCTGAGGCTGGGGCCAACTGGTG 3180
QY 3181 TAAATCTGTGGCTGTGCAAGGATTCAGAGCAGACCTGCAATCTTGTGACGCTGGGGA 3240
DB 3181 TAAATCTGTGGCTGTGCAAGGATTCAGAGCAGACCTGCAATCTTGTGACGCTGGGGA 3240
QY 3241 GGGGGAAGGAGCAACCCCTTATTAACCTTTTCTCTCAGGCCCAAGATTAACACT 3300
DB 3241 GGGGGAAGGAGCAACCCCTTATTAACCTTTTCTCTCAGGCCCAAGATTAACACT 3300
QY 3301 CTGGCTTCCCTTCCACTCCATCAGAGTGAAGGTTGCAAGAGGAGTTAAAAA 3360

Db	3301	CTGGCCCTTCCCCCTCCACCTCCCATCAGAGTGGAGGGTTGCGAGGGGATAAAA	3360
Qy	3361	CTTACATGTCMAACATCATGAGTCACAGATATATGATCAGTATGTTAGAGCAAAAA	3420
Db	3361	CCTACATGTCMAACATCATGAGTCACAGATATATGATCAGTATGTTAGAGGCAAAAA	3420
Qy	3421	GGAAATCTGCAGGGCTTAACTGGGTTAAATGCTGTAAGTCTGTGCGACATGTTGTGTCTG	3480
Db	3421	GGAAATCTGCAGGGCTTAACTGGGTTAAATGCTGTAAGTCTGTGCGACATGTTGTGTCTG	3480
Qy	3481	ACTGAAAAACGGGCACTGAGTGTGCAGCTGTTCAAGTCTGTGCGTAGGTTACAGACTGCA	3540
Db	3481	ACTGAAAAACGGGCACTGAGTGTGCAGCTGTTCAAGTCTGTGCGTAGGTTACAGACTGCA	3540
Qy	3541	GGTTTGTGTGTAATTGCCCCAAGGCAAGTGGGTGAATCCCTTCCATGTTTTAAAGAT	3600
Db	3541	GGTTTGTGTGTAATTGCCCCAAGGCAAGTGGGTGAATCCCTTCCATGTTTTAAAGAT	3600
Qy	3601	TGATATGATGGCTGCATCTCAAGGACCATGGAATAATGAATGACACTTATATGTCT	3660
Db	3601	TGATATGATGGCTGCATCTCAAGGACCATGGAATAATGAATGACACTTATATGTCT	3660
Qy	3661	CTAAGCTAAGTGAAGCAAGTCTTTTGAGAGACACTGTCTAGAGATGTGGCAACAGAC	3720
Db	3661	CTAAGCTAAGTGAAGCAAGTCTTTTGAGAGACACTGTCTAGAGATGTGGCAACAGAC	3720
Qy	3721	TACAGACAGTATCTGTACAGATTAAGAGACAGAGGGGGTGTAGAAATTCCTCTTA	3780
Db	3721	TACAGACAGTATCTGTACAGATTAAGAGACAGAGGGGGTGTAGAAATTCCTCTTA	3780
Qy	3781	TCAAAAGGAAACTGAGTCTGTGCACCTGCAAAAGTGAATGCTTCCCTAGACATCATGACTT	3840
Db	3781	TCAAAAGGAAACTGAGTCTGTGCACCTGCAAAAGTGAATGCTTCCCTAGACATCATGACTT	3840
Qy	3841	TGTCCTTGGGAGCCAGCATCTGTGAACTTCAAGTCTGAGAGATGAGAGGCTCCCTCA	3900
Db	3841	TGTCCTTGGGAGCCAGCATCTGTGAACTTCAAGTCTGAGAGATGAGAGGCTCCCTCA	3900
Qy	3901	GCTGAAGCTATGCGATAGCCAGGGTTGAAAGGGGGAAAGGAGAGCCTGGGATGGGAGC	3960
Db	3901	GCTGAAGCTATGCGATAGCCAGGGTTGAAAGGGGGAAAGGAGAGCCTGGGATGGGAGC	3960
Qy	3961	TTGTGTGTTGAGGCGAGGGGACAGATATTAAAGCTTGAAGAGAAAGTGACCTTACCAG	4020
Db	3961	TTGTGTGTTGAGGCGAGGGGACAGATATTAAAGCTTGAAGAGAAAGTGACCTTACCAG	4020
Qy	4021	TTGTTCAACTCAACCTTCAAGTTAAAAATACTAGAGTAAAGGGCTGGGTATGGGGAAGTG	4080
Db	4021	TTGTTCAACTCAACCTTCAAGTTAAAAATACTAGAGTAAAGGGCTGGGTATGGGGAAGTG	4080
Qy	4081	GTTGTAGACGCTCTGTCTCTCTCTGTGCATGCCCTGAGGCCCTTTGGGAGAGAAATGT	4140
Db	4081	GTTGTAGACGCTCTGTCTCTCTCTGTGCATGCCCTTGGGGAAGAAATGT	4140
Qy	4141	GCCCAAGGACTTAAAAAAGGCCATGAGACCGAGAGGGGAGAGGCAACAGACTTTCATGG	4200
Db	4141	GCCCAAGGACTTAAAAAAGGCCATGAGACCGAGAGGGGAGAGGCAACAGACTTTCATGG	4200
Qy	4201	GCAAACTTTGGGGGCCCTGATGATGATGCAAGAACTCCCACTCGATCCCTTCTTC	4260
Db	4201	GCAAACTTTGGGGGCCCTGATGATGATGCAAGAACTCCCACTCGATCCCTTCTTC	4260
Qy	4261	TTCTTAACGACAGAGGGAATCTCGAGTTTACCACTCCCTATCAGTATAGAAAAAGTGA	4320
Db	4261	GGAGCCAGAGGACGAGGGAAGTGGAGGAGAG-----	4321
Qy	4321	AAGTGAATTTACCACTCCCTATCAGTATAGAGAAAGTGAAGTCAAGTTTACCACTC	4380
Db	4322	-----	4321
Qy	4381	CCTATCAGTATAGAAAAAGTGAAGTCAAGTTTACCACTCCCTATCAGTATAGAGAA	4440

D	b	4292	-----	4291
O	y	4441	AAGTAAAGTCGAGTTTACCACTCCCTATCAGGTATAGAGAAAAGTGAAGTCGAGTTTA	4500
D	b	4292	-----	4291
O	y	4501	CCACTCCCTATCAGTGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTAT	4560
D	b	4292	-----	4291
O	y	4561	AGAGAAAAGTGAAGTCGAGCTCGGTACGACGAGGACTCCAAATTTTAGCAGCAGGCA	4620
D	b	4292	-----GGTCCAGACAGAGACTCCAAATTTAGCAGCAGGCA	4328
O	y	4621	TATGGATGGATATATAAGGGCTGAGACCTGAGAGCTGTCAGAGATTTCTCCAGCCCA	4680
D	b	4329	TATGGATGGATATATAAGGGCTGAGACCTGAGAGCTGTCAGAGATTTCTCCAGCCCA	4388
O	y	4681	GGTAAAGGAGATTTGGGTGGGGGCTCTTCCACCACACAGACTCTCCCACTAGAA	4740
D	b	4389	GGTAAAGGAGATTTGGGTGGGGGCTCTTCCACCACAGACTCTCTCCCACTAGAA	4448
O	y	4741	GGAAACTGCTTTCTCTGGAAGTGGGCTTCAAGCCCGGTACAGATTTTGA CAGGTGGCTT	4800
D	b	4449	GGAAACTGCTTTCTCTGGAAGTGGGCTTCAAGCCCGGTACAGATTTTGA CAGGTGGCTT	4508
O	y	4801	CCAACAGCCTGGGAATTTCTCAGTGGACAGAGATTTTCCAAAGAAAACCTGATGCCCT	4860
D	b	4509	CCAACAGCCTGGGAATTTCTCAGTGGACAGAGATTTTCCAAAGAAAACCTGATGCCCT	4568
O	y	4861	TCCCTTACGCTGTCTTCATCTTCTCTGGGAGTGTCTCTCTCCGCTCTTGGTTATC	4920
D	b	4569	TCCCTTACGCTGTCTTCATCTTCTCTGGGAGTGTCTCTCTCCGCTCTTGGTTATC	4628
O	y	4921	TTGGCTTCGTGCTTACGAGAAATTTGGCCGTGTGTCACACTGCATCTTCTCTACGT	4980
D	b	4629	TTGGCTTCGTGCTTACGAGAAATTTGGCCGTGTGTCACACTGCATCTTCTCTACGT	4688
O	y	4981	CTCCGTGCTTGGCTTGCCTTCTTGGCGTTCCTTCTTCCACCCTATTTCTACCTTACC	5040
D	b	4689	CTCCGTGCTTGGCTTGCCTTCTTGGCGTTCCTTCTTCCACCCTATTTCTACCTTACC	4748
O	y	5041	TTTTCTCCCTCTTCAATTTGATATCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	5100
D	b	4749	TTTTCTCCCTCTTCAATTTGATATCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	4808
O	y	5101	CTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	5160
D	b	4809	CTTCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	4868
O	y	5161	CTGTGTACAGAGTGTGAAGATCACACTGGGTTTCCACCCTTATGTAAACAATCTTCCA	5220
D	b	4869	CTGTGTACAGAGTGTGAAGATCACACTGGGTTTCCACCCTTATGTAAACAATCTTCCA	4928
O	y	5221	GTAGAGCAGAGCTTCAAGTGTGTGGGATCTCTTAACTCTTCAACCCCGAGCTTATC	5280
D	b	4929	GTAGAGCAGAGCTTCAAGTGTGTGGGATCTCTTAACTCTTCAACCCCGAGCTTATC	4988
O	y	5281	CTGTTCATCTGTGAGATCTTATAGATTTGATCTTCCAGCCTCTGTACTCTTCTTCT	5340
D	b	4989	CTGTTCATCTGTGAGATCTTATAGATTTGATCTTCCAGCCTCTGTACTCTTCTTCT	5048
O	y	5341	GCTGTTCCTCTCTCTGTCTGTCAGAGTGGCCACTGTGTGTCTGTCTTCCAGCTGTGTCCAC	5400
D	b	5049	GCTGTTCCTCTCTCTGTCTGTCAGAGTGGCCACTGTGTGTCTGTCTTCCAGCTGTGTCCAC	5108
O	y	5401	ATTCTTCAGAGATTTCTGAAAAAATTTAACAGAGTGAAGATTTTCCCTGTAGACAGCAGA	5460
D	b	5109	ATTCTTCAGAGATTTCTGAAAAAATTTAACAGAGTGAAGATTTTCCCTGTAGACAGCAGA	5168
O	y	5461	TACAGATTTCTCCGGAAGTCAGGCTTCCAGCCTCTCTTCTCTGTGCCAGTGTCCGGCA	5520
D	b	5169	TACAGATTTCTCCGGAAGTCAGGCTTCCAGCCTCTCTTCTCTGTGCCAGTGTCCGGCA	5228

QY 5521 CTCTTAGCAAACTCAGGCAACCTTACCCACATAGACCTTGACAGAGAAGACGACCT 5580
DB 5229 CTCTTAGCAAACTCAGGCAACCTTACCCACATAGACCTTGACAGAGAAGACGACCT 5288
QY 5581 TTACATGAGATCTCTGCTGAGAGAGCCATAGAGCTACGGTGTAAAAAGAGCAGGAAGTGT 5640
DB 5289 TTACATGAGATCTCTGCTGAGAGAGCCATAGAGCTACGGTGTAAAAAGAGCAGGAAGTGT 5348
QY 5641 GGTGTAGAAAGTCAGAGACTTTCATATGAAAGCTTACGCCACACAGAAATGACAGACAGA 5700
DB 5349 GGTGTAGAAAGTCAGAGACTTTCATATGAAAGCTTACGCCACACAGAAATGACAGACAGA 5408
QY 5701 TCCCTCTATCTCCCCCATAGAGTTGAGTCGAC 5735
DB 5409 TCCCTCTATCTCCCCCATAGAGTTGAGTCGAC 5443

RESULT 5

US-10-027-655-5
; Sequence 5, Application US/10027655
; Publication No. US20030204206A1
; GENERAL INFORMATION:
; APPLICANT: Medtronic, Inc.
; APPLICANT: Padua, Rodolfo
; APPLICANT: Schu, Carl
; APPLICANT: Bonner, Matthew
; APPLICANT: Donovan, Maura
; APPLICANT: Soykan, Orhan
; TITLE OF INVENTION: Electrically Responsive Promoter System
; FILE REFERENCE: P9406.00
; CURRENT APPLICATION NUMBER: US/10/027,655
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 5
; LENGTH: 1679
; TYPE: DNA
; ORGANISM: Mouse
US-10-027-655-5

Query Match 22.3%; Score 1279.8; DB 17; Length 1679;

Best Local Similarity 82.2%; Pred. No. 0;

Matches 1616; Conservative 0; Mismatches 57; Indels 292; Gaps 1;

QY 3767 GAATTCCTTACTATCAAAAGGAAAGTGTGCACTGCAAAAGTGTCTTCCCT 3826
DB 1 GAATTCCTTACTATCAAAAGGAAAGTGTGCACTGCAAAAGTGTCTTCCCT 60
QY 3827 AGACATCATATCTTGTCTCTGGGAGCCAGCACTGTGAACTTCAGGTCTGAGAGATTA 3886
DB 61 AGACATCATATCTTGTCTCTGGGAGCCAGCACTGTGAACTTCAGGTCTGAGAGATTA 120
QY 3887 GAGAGCTCCCTCAGCTGAAAGCTATGACATAGCCAGGTTGAAAAGGGGAAAGGAG 3946
DB 121 GAGAGCTCCCTCAGCTGAAAGCTATGACATAGCCAGGTTGAAAAGGGGAAAGGAG 180
QY 3947 CCTGGATGGAGCTTGTGTGTGAGGAGGAGGAGATATTAAGCTTGAAGAGAGG 4006
DB 181 CCTGGATGGAGCTTGTGTGTGAGGAGGAGGAGATATTAAGCTTGAAGAGAGG 240
QY 4007 TGACCTTACCCAGTGTGTCACTCACCTTCAGATTAATAAATACTGAGGTAAAGGCTT 4066
DB 241 TGACCTTACCCAGTGTGTGTCACTCACCTTCAGATTAATAAATACTGAGGTAAAGGCTT 300
QY 4067 GGGTAGGGAGGTGTGTGAGAGGCTCTGTCTCTCTGCAATGACCTGAGGCTCTTTG 4126
DB 301 GGGTAGGGAGGTGTGTGAGAGGCTCTGTCTCTCTCTATGTGCCCATGAGGCTTTG 360
QY 4127 GGGAGAGGAATGTGCCAAGACTAAAAAGGCAATGAGCCAGAGGGGCAAGGCA 4186
DB 361 GGGAGAGGAATGTGCCAAGACTAAAAAGGCAATGAGCCAGAGGGGCAAGGCA 420

QY 4187 CAGACCTTTATGGGAAACCTTGGGGCCCGATGATGATGATTCACAAAGACTCGCAAT 4246
DB 421 CAGACCTTTATGGGAAACCTTGGGGCCCGATGATGATGATGATTCACAAAGACTCGCAAT 480
QY 4247 CGATACCTTCTCTTCTTCAACGAGAGAGGAGAACTCGAGTTTACCATCTCCATCAGTG 4306
DB 481 GGATCAAAAGAGAGAGAGCAGAGACAGAGGAGGAGTGGAGGAG----- 525
QY 4307 ATTAGAAAAAGTGAAGTGAATTTACATCTCCCTATCAGTATAGAAAAAGTGAAGT 4366
DB 526 ----- 525
QY 4367 CGAGTTTACACCTCCCTATCAGTATAGAAAAAGTGAAGTGAATTTACATCTCCCTA 4426
DB 526 ----- 525
QY 4427 TCAGTATAGAAAAAGTGAAGTGAATTTACATCTCCCTATCAGTATAGAAAAAGT 4486
DB 526 ----- 525
QY 4487 GAAAGTCAGTTTACACCTCCCTATCAGTATAGAAAAAGTGAAGTGAATTTACATCT 4546
DB 526 ----- 525
QY 4547 TCCCTATCAGTATAGAAAAAGTGAAGTGAATTTACATCTCCCTATCAGTATAGAAAA 4606
DB 526 -----GGTCCAGCAGAGACTCCAAAT 548
QY 4607 TTAGCAGAGCAGATATGAGATGAGATATTAAGGGGCTGAGACACTGAGAGCTGTGAG 4666
DB 549 TTAGCAGAGCAGATATGAGATGAGATATTAAGGGGCTGAGACACTGAGAGCTGTGAG 608
QY 4667 ATTTCTCCAAACCAAGTAAAGAGAGTTCGGGTGGGGCTTTCAACCAACAGACCT 4726
DB 609 ATTTCTCCAAACCAAGTAAAGAGAGTTCGGGTGGGGCTTTCAACCAACAGACCT 668
QY 4727 CTCCTCCATAGAGAAACCTCTTCTGTAAGTGGGGTTCAGGCGGTGAGAGATCT 4786
DB 669 CTCCTCCATAGAGAAACCTCTTCTGTAAGTGGGGTTCAGGCGGTGAGAGATCT 728
QY 4787 GACAGGATGAGCTTCACAGAGCTGAGAAAGTTCAGTGGCAGAGAGTTTCCAAAGAA 4846
DB 729 GACAGGATGAGCTTCACAGAGCTGAGAAAGTTCAGTGGCAGAGAGTTTCCAAAGAA 788
QY 4847 CACTGATGAGCTTCCTTCACTGTCTTCTTCATCTTCTCTGGGATGCTCTCC 4906
DB 789 CACTGATGAGCTTCCTTCACTGTCTTCTTCATCTTCTCTGGGATGCTCTCC 848
QY 4907 CGTCTTGTATATCTTGTGCTTCTTCTGACAGAAATTTGACCTGTGCTGCTGCA 4966
DB 849 CGTCTTGTATATCTTGTGCTTCTTCTGACAGAAATTTGACCTGTGCTGCTGCA 908
QY 4967 TCTTCTCTACGTCTGCGTGCCTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5026
DB 909 TCTTCTCTACGTCTGCGTGCCTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 968
QY 5027 TTTCTCACTTCACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5086
DB 1029 TCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1088
QY 5087 TCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5146
DB 1089 TCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1148
QY 5207 TAAACATTTTCAAGTGAAGCCAGAGCTTCAAGTGTGCTGCTCTTACCTTCTCTA 5266
DB 1149 TAAACATTTTCAAGTGAAGCCAGAGCTTCAAGTGTGCTGCTCTTACCTTCTCTA 1208
QY 5267 CCCCCTGCTTGTCTGTTTCACTCTGCTGATCAGATCTGATATGATGCTTCCAGCTCTG 5326


```

Db 3530 CATCATGTACACAAAAGAAAGTAAAGAAATCTGCAGCTTGATGGGTAAAGTGA 3589
Qy 3453 AAAGTCTGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3512
Db 3590 AAATCTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3633
Qy 3513 GTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3572
Db 3634 --TCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3691
Qy 3573 GTGAATCTCTCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3632
Db 3692 GAGATACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3750
Qy 3633 AATTAAGATGACATCTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3689
Db 3751 GGAAGGGGTGGGACCTCCAGACGTGTCTCCAAACAGGGGTGGCTCCAGACCTTGGGA 3810
Qy 3690 ACACGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3749
Db 3811 ACATCTGTCTGAAGCTTGGGAAACG-----AAGAGACCAAGGCAATGGCACTTATGCA 3865
Qy 3750 AGAGAGAGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3809
Db 3866 ACTGAGGCCAGACAGAA-----TTTCTGACAAAAGAAACCTGAGCCATGAGATGGA 3919
Qy 3810 AAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3869
Db 3920 CAACAGATCCCTTCCCTGGGCA--CAATCTGACGCTTTTGTGTGTGTGTGTGTGTGTGT 3977
Qy 3870 TCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3929
Db 3978 CAGTACTTAACAGACAGAAAGTGTCTCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGT 4037
Qy 3930 AAAAG-GGAGAGGAGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3988
Db 4038 AAGAGAGAGGCTGGGAAACAGGGAGAGAAAGCCCATGTGTGTGTGTGTGTGTGTGTGTGT 4097
Qy 3989 TAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4048
Db 4098 TTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4156
Qy 4049 TAACTGAGTAAAGG-----CTGGGTAGGGAGGTGTGTGTGTGTGTGTGTGTGTGTGT 4103
Db 4157 TAACTGAGTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4216
Qy 4104 TC-TGCATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4160
Db 4217 ACATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4276
Qy 4161 CCATGAGCCAGAGGGGCGAGGGGCAACAGCTTTTCAATGGGCAACCTTGGGGCCC 4216
Db 4277 CCGTGTGAGCCAGAGAGGCTGGGGCAGACAGCTTTTCTTGGGCAAAATCAGGGGGCCC 4332

```

RESULT 7

```

US-10-798-037-1
; Sequence 1, Application US/10798037
; Publication No. US20040229360A1
; GENERAL INFORMATION:
; APPLICANT: Robbins, Jeffrey
; APPLICANT: Children's Hospital Medical Center
; TITLE OF INVENTION: Cardiac-preferred Genetic Alteration of
; TITLE OF INVENTION: Transgenic Rabbit
; FILE REFERENCE: CHM02/GN054
; CURRENT APPLICATION NUMBER: US/10/798,037
; CURRENT FILING DATE: 2004-03-11
; PRIOR APPLICATION NUMBER: 60/454,947
; PRIOR FILING DATE: 2003-03-13
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 5190

```

```

; TYPE: DNA
; ORGANISM: Oryzotylagus cuniculus
US-10-798-037-1
Query Match 7.1%; Score 407.6; DB 20; Length 5190;
Best Local Similarity 58.8%; Pred. No. 1,6e-114;
Matches 1265; Conservative 0; Mismatches 699; Indels 188; Gaps 25;

Qy 2099 ACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2157
Db 2140 ACCCTTGGGCAAGATATCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2199
Qy 2158 GCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2201
Db 2200 AGCTGATTTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2259
Qy 2202 GAGACAGGGGGCTTAGAAGAAAGAGGCAACCAAGGCCACACAGAGGGCAGAG 2261
Db 2260 CAGAGAGGGGGCCACAGGAAAG--GAAAGGGCAAAACAGGCCAGACCG--GCCCGAG 2314
Qy 2262 CCCAGACTGAGTTAATCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2321
Db 2315 CCCAGAGCCAGGCTAACCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2372
Qy 2322 TGACCAACCATCCCATGAGCCCATTAACCATTAAGTTGGCTGTGTGTGTGTGTGTGT 2381
Db 2373 TGCCCAACCTTAGGCCCGCCCTTACCACTCTCCAAAGTTGTGTGTGTGTGTGTGTGT 2432
Qy 2382 AGGTTT-CCTGAGAGACAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2440
Db 2433 GGGTCTCCGAGGACAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2492
Qy 2441 TGTCTCACTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2500
Db 2493 TATTC-CAGCCCTTATATATATATATATATATATATATATATATATATATATATAT 2551
Qy 2501 CATGGGCTAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2560
Db 2552 CAGGGATTAACC-CGAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2602
Qy 2561 CAGACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2617
Db 2603 GAGACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2662
Qy 2618 GGAATGCAAGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2677
Db 2663 GGAATGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2722
Qy 2678 GAGAGAGGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2737
Db 2723 GAGAGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2774
Qy 2738 TGGGCACTTTGGGAGACACCA--CATTAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2794
Db 2775 TGACAAACCTTGGGGGAAACCCATCTTCAAGGCAAGTGTGTGTGTGTGTGTGTGTGTGT 2834
Qy 2795 GCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2854
Db 2835 CCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2894
Qy 2855 GTGGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2914
Db 2895 G-GTGGAGGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2953
Qy 2915 AGGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2974
Db 2954 AGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3013
Qy 2975 GCTAAAGACCAACATTAAGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3032
Db 3014 GTTGAAGACCAACATTAAGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3073
Qy 3033 GTCTGTGAGCTTGGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3092

```

Db 3074 ACCCTGAGGAGCTGGCTCTTGAAGACCCCTCCCAAGACAGGCAATGAGAGTGTAGC 3133
Qy 3093 CCCAGAGAGGCTTAGAGAGGAGGCACTTGGAGAGGGGTCAGTCTGCAAGAGCCCTATGCAT 3152
Db 3134 TCAGAGAGGAGATGACAGGCTTCTATCTGGAGAGGGGCAAGCTTACAGAGCCCTCCCTCC 3193
Qy 3153 GGAATCTGAGAGCCTGGGGCCCAACTGTGTAAATCTCTGGGCTGGCCAGGATTCMAAGCA 3212
Db 3194 CAACCCCTGCAACCCCTGTCAGAGCCCTGTGGAACCGAGAGCCCAATTTGGTGTAAAT 3253
Qy 3213 GCACCTGCAATCCTTGGCAGGCTTGGGAGGCGAGAGGAGCAACCCCTTATACCT 3272
Db 3254 CCCAGAGCCTGTCTGGGCGCCCAACCATCTCGGACACCTGCAACCTCTGGCAGCTG 3313
Qy 3273 TTCTCCCTCAGCCCCAGAGATTAAACACTTGGACCTTCCCTCCCACTCCATCAGCA 3332
Db 3314 GAGAGGCGAGAGGAGAGGAGGCCCCCAACCCCTCCTCTCAGGCGGGGATTAAT 3373
Qy 3333 GTGAGAGGTTGCAAGAGGAGGATTAACCTACATCTCAACATCATGTGACATAT 3392
Db 3374 GCAGAGGTGGCAGACAGAGGATTAAGCCTTGGCGCA-----GATGT 3418
Qy 3393 ATGATCAATATGTGTAGAGGCAAGAAATCTCAGAGCTTAACCTGGTTAATGTGT 3452
Db 3419 GAGCATCGGCTGTGGAAGACGAGGAGATCTGTAAAGCTGAAGTGTGAAA----- 3472
Qy 3453 AAAGTGTGTGATGTGTGTGTGTCTGACGTAAGAAAGGGGAGCTGTGACAGCTTTCA 3512
Db 3473 -----TCTGATGTGTGTGTGTCTCA 3492
Qy 3513 GTTCTGTGCTGAGGTTACAGACTGAGGTTGTGTGTAAATTTGCCAAGGCAAGTGTG 3572
Db 3493 AGTCTGACGCGAGGAGGCTGTGACCAAGGCTGTGCGCAGGTGTGCTGAGGAGGCTG 3552
Qy 3573 GTGATCCCTTCCATGCTTTAAAGAGATGATGATGCTGCTGCTCAAGGACATGGA 3632
Db 3553 GAGAAATCTTCCACTGGGTGATGGGTTTGCATAGGGCCCACTGTATGGGAGGACCGGG 3612
Qy 3633 AATAGAGATGAGCACTATATGTGTCTCTAAGCTAAGTAGCAAGGCTTTTGAAGACA 3692
Db 3613 GAAATGCTGGGGCTCCAGATGCAATTCACAGCAAGAGCCCTTATAGTCTCTGTGACACA 3672
Qy 3693 CTTGTAGAGATGTGGGCAACAGAGCTACAGACATATCTGTACAGAGTAGAGAGA 3752
Db 3673 CCTGCC-----TGGAAATAGAAGCCACAGGAGTGTGTGAGAGATTTGGGAGAGG 3724
Qy 3753 GAGAGGGGGGTAGAAATCTCTTACTATCAAGAGGAACTGAGTCTGTGACCTTGCAAG 3812
Db 3725 ACAAG-----AGGAAACTGAGTCTGAGCTGGCTAG 3757
Qy 3813 TGGATGCTCTCCCTAGACATGATCTTGTCTGTGGGAGAGCAGCACTGTGAACTTCA 3872
Db 3758 TAGACGCTGTCTCAGCTCACAG-----TCCA 3785
Qy 3873 GGTCTGAGAGAGTAGAGGCTCCCTCAGCT--GAACTATGAGATAGCCAGGTTGAA 3931
Db 3786 GGTCTGAGAGAGTAGAGGACATTTGAGAGCTGTGGCTGTGTGAAGACAGGCTGGGTAAA 3845
Qy 3932 AGGGGGAAGGGA--GAGCTTGGAGATGGAGCTTGTGTGTGAGGACAGGAGACAGATTTA 3990
Db 3846 TGGGAGAGGAGTAGAGGACAAAGGAGGAACTGAGCTGGGGGCAAGGAGCAAGCATTC 3905
Qy 3991 AGCTGGAAGAGAGTAGGACCTTACCAAGTTGT--TCAATCAACCTTTCAGATTTAAAA 4048
Db 3906 GTCTTATATGAAAAAGTAGACCTTCAACCCAGTGTCTCACTCACTTTCAGGTTAAAAA 3965
Qy 4049 TAACTAGAGTAAGGAGCTTGGTGAAGGAGATGTGT-----GAGAGCTCTCTGTCTCT 4104
Db 3966 TAACTAGAGTAAGGAGCTTGGTGAAGGAGATGTGTGTGAAGAGATCTGTCTTCCA 4025
Qy 4105 CTGATGCTCTGAGGAGCTTTTGGGAGAGAGATGTGTGCAAGAGCTTAAAAAGGCGCAT 4164

Db 4026 CTATCTGCCATCAGGCTCTGTGAGGGGCGGAATGTGCTCAAGACTAAAAAGGCCCT 4085
Qy 4165 GGAGCCAGAGGGGCGAGGCAACAGACCTTTCAATGAGCAAACTTGGGCGCC 4216
Db 4086 GAGGCCGAGGGGCTGGGCGACGACCTTTCAATGAGCAAAATGTGGGGGCC 4137

RESULT 8
US-09-874-389-7
Sequence 7, Application US/09874389
Patent No. US20020152489A1
GENERAL INFORMATION:
APPLICANT: Bujard, Hermann
Goessen, Manfred
TITLE OF INVENTION: Animal Transgenic for a Tetracycline- Inducible
Transcription
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/874,389
FILING DATE: 26-Dec-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/161,902
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-1994
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-94
APPLICATION NUMBER: US 08/260,452
FILING DATE: 14-JUNE-1994
APPLICATION NUMBER: US 08/076,327
FILING DATE: 14-JUNE-1993
APPLICATION NUMBER: US 08/076,726
FILING DATE: 14-JUNE-1993
ATTORNEY/AGENT INFORMATION:
NAME: Decont, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BH1-009CP3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-874-389-7

Query Match 5.4%; Score 311.8; DB 9; Length 520;
Best Local Similarity 99.4%; Pred. No. 2,7e-85;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 4277 GGAATCGAATTACCACTCCCTATCATGATAGAGAAAAAGTGAAGTCAATTACAC 4336
Db 63 GGAATCGAATTACCACTCCCTATCATGATAGAGAAAAAGTGAAGTCAATTACAC 122
Qy 4337 TCCCTATCATGATAGAGAAAAAGTGAAGTCAATTACCACTCCCTATCATGATAGAG 4396

DB 123 TCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCATCCCTATCAGTATAGAG 182
QY 4397 AAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCAGTT 4456
DB 183 AAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCAGTT 242
QY 4457 TACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTG 4516
DB 243 TACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTG 302
QY 4517 ATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGT 4576
DB 303 ATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGT 362
QY 4577 CGAGCTCGGTACCG 4591
DB 363 CGAGCTCGGTACCG 377

RESULT 9
US-09-921-650-7
Sequence 7, Application US/09921650
Publication No. US2003022315A1
GENERAL INFORMATION:
APPLICANT: Buijard, Hermann
Goossen, Manfred
TITLE OF INVENTION: Tetracycline-Inducible Transcriptional
Activator Fusion Proteins
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/921,650
FILING DATE: 03-Aug-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/912,650
FILING DATE: 2001-08-03
APPLICATION NUMBER: US 08/485,978
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/383,754
FILING DATE: 03-FEB-1995
APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-1994
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-1994
APPLICATION NUMBER: US 08/260,452
FILING DATE: 14-JUNE-1994
APPLICATION NUMBER: US 08/076,327
FILING DATE: 14-JUNE-1993
APPLICATION NUMBER: US 08/076,726
FILING DATE: 14-JUNE-1993
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-0096CNDV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-921-650-7
Query Match
Best Local Similarity 99.4%; Score 311.8; DB 10; Length 520;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4277 GGAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCA 4336
DB 63 GGATCTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCA 122
QY 4337 TCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAGAG 4396
DB 123 TCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAGAG 182
QY 4397 AAAAGTGAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTT 4456
DB 183 AAAAGTGAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTT 242
QY 4457 TACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTG 4516
DB 243 TACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTG 302
QY 4517 ATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGT 4576
DB 303 ATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGT 362
QY 4577 CGAGCTCGGTACCG 4591
DB 363 CGAGCTCGGTACCG 377

RESULT 10
US-09-241-347-7
Sequence 7, Application US/09241347
Publication No. US20040003417A1
GENERAL INFORMATION:
APPLICANT: Buijard, Hermann
Goossen, Manfred
TITLE OF INVENTION: Animals Transgenic for a Tetracycline-
Controlled Transc
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/241,347
FILING DATE: 02-Feb-1999
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/486,814
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/383,754
FILING DATE: 03-FEB-1995
APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-1994
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-1994
APPLICATION NUMBER: US 08/260,452
FILING DATE: 14-JUNE-1994
APPLICATION NUMBER: US 08/076,327
FILING DATE: 14-JUNE-1993

```
Query Match      5.4%; Score 311.8; DB 11; Length 520;
Best Local Similarity 99.4%; Pred. No. 2.7e-85;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

Qy	4277	GGAATCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCAGATTACAC	4336
Db	63	GGATCTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCAGATTACAC	122
Qy	4337	TCCCTATCAGTATAGAGAAAAGTGAAGTCAGATTACCACTCCCTATCAGTATAGAG	4396
Db	123	TCCCTATCAGTATAGAGAAAAGTGAAGTCAGATTACCACTCCCTATCAGTATAGAG	182
Qy	4397	AAAAGTGAAGTCAGATTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCAGATT	4456
Db	183	AAAAGTGAAGTCAGATTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCAGATT	242
Qy	4457	TACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCAGATTACCACTCCCTATCAGTG	4516
Db	243	TACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCAGATTACCACTCCCTATCAGTG	302
Qy	4517	ATAGAGAAAAGTGAAGTCAGATTACCACTCCCTATCAGTATAGAGAAAAGTGAAGT	4576
Db	303	ATAGAGAAAAGTGAAGTCAGATTACCACTCCCTATCAGTATAGAGAAAAGTGAAGT	362
Qy	4577	CGAGCTCGTACGAG 4591	
Db	363	CGAGCTCGTACCGG 377	

```

RESULT 11
US-10-134-643-17
; Sequence 17, Application US/10134643
; Publication No. US20030113898A1
; GENERAL INFORMATION:
; APPLICANT: OLSEN, JOHN C.
; APPLICANT: MITROPHANOUS, KYRIACOS ANDREOU
; APPLICANT: ROHILL, JONATHAN
; APPLICANT: KINGSMAN, ALAN JOHN
; APPLICANT: ELIARD, FLONA MARGARET
; TITLE OF INVENTION: METHODS FOR PRODUCING HIGH TITRE VECTORS AND
; TITLE OF INVENTION: COMPOSITIONS USED IN SUCH METHODS
; FILE REFERENCE: 078883-0148
; CURRENT APPLICATION NUMBER: US/10/134,643
; CURRENT FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: 60/287,048
; PRIOR FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 990
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

```

Query Match	5.4%;	Score 311.4;	DB 15;	Length 950;
Best Local Similarity	-99.7%;	Pred. No. 5.1e-85;		
Matches 312; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

OY	4279	AACTCGAGTTTACCACTCCCTATCAGTGATGAGAAAAGTGAAGTTCGAGTTTACCACTC	4338
Db	156	AACTCGAGTTTACCACTCCCTATCAGTGATGAGAAAAGTGAAGTTCGAGTTTACCACTC	215
OY	4339	CCATCAGTGATGAGAAAAGTGAAGTTCGAGTTTACCACTCCCTATCAGTGATGAGAAA	4398
Db	216	CCATCAGTGATGAGAAAAGTGAAGTTCGAGTTTACCACTCCCTATCAGTGATGAGAAA	275
OY	4399	AAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAAAAGTGAAGTTCGAGTTT	4458
Db	276	AAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAAAAGTGAAGTTCGAGTTT	335
OY	4459	CCACTCCCTATCAGTGATGAGAAAAGTGAAGTTCGAGTTTACCACTCCCTATCAGTGAT	4518
Db	336	CCACTCCCTATCAGTGATGAGAAAAGTGAAGTTCGAGTTTACCACTCCCTATCAGTGAT	395
OY	4519	AGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAAAAGTGAAGTTCG	4578
Db	396	AGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTGATGAGAAAAGTGAAGTTCG	455
OY	4579	AGCTTCGGTACCG	4591
Db	456	AGCTTCGGTACCG	468

```

RESULT 12
US-10-102-143-12/c
/ Sequence 12. Application US/10102143
/ Publication NO. US20030185851A1
/ GENERAL INFORMATION:
/ APPLICANT: Solidati, Dominique
/ APPLICANT: Meissner, Markus
/ TITLE OF INVENTION: Text transactivator system
/ FILE REFERENCE: 04630/016001
/ CURRENT APPLICATION NUMBER: US/10/102.143
/ CURRENT FILING DATE: 2003-01-21
/ NUMBER OF SEQ ID NOS: 31
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 12
/ LENGTH: 4438
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: (1193)..(1885)
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence
/ OTHER INFORMATION: ptefo/Sagl-HXGPR
US-10-102-143-12

```

Query Match	Similarity	Score	DB	Length
Best Local	5.4%	96.9%	311	4438
Matches	Conservative	Pred. No.	1.5e-84	
317	0	Mismatches	10	Indels
		Gaps	0	
QY	4263	CTAAACGACAGAGAGGAAGTTCAGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAA	43322	
DB	979	CTTATCCATACCGTCGACCTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAA	920	
QY	4333	GTCCAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAAAGTCGAGTTTACCACTCCC	4388	
DB	919	GTCCAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAAAGTCGAGTTTACCACTCCC	860	
QY	4383	TATCAGTATAGAGAAAAGTGAAAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAAA	44447	
DB	859	TATCAGTATAGAGAAAAGTGAAAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAAA	800	

```
QY 4443 GTGAAAGTCAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACC 4502
DB 799 GTGAAAGTCAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACC 740
QY 4503 ACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAG 4562
DB 739 ACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAG 680
QY 4563 AGAAAAGTGAAGTCGAGTCGCTCGGTACC 4589
DB 679 AGAAAAGTGAAGTCGAGTCGCTCGGTACC 653
```

RESULT 13

```
US-10-102-143-10/c
; Sequence 10, Application US/10102143
; Publication No. US20030185851A1
; GENERAL INFORMATION:
; APPLICANT: Soldati, Dominique
; APPLICANT: Meisner, Markus
; TITLE OF INVENTION: TET transactivator system
; FILE REFERENCE: 04630/016001
; CURRENT APPLICATION NUMBER: US/10/102,143
; CURRENT FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 10
; LENGTH: 4479
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1193)..(1924)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1102)
; OTHER INFORMATION: n is disclosed as an asterisk
; OTHER INFORMATION: Description of Artificial Sequence: pTetO7Sag1-GFP
US-10-102-143-10
```

```
Query Match 5.4%; Score 311; DB 16; Length 4479;
Best Local Similarity 96.9%; Pred. No. 1.5e-84;
Matches 317; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 4263 CTAAACGACAGGAGGAACTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGA 4322
DB 979 CTATCGATACCGTCGACCTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGA 920
QY 4323 GTGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCC 4382
DB 919 GTGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCC 860
QY 4383 TATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAAA 4442
DB 859 TATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAAA 800
QY 4443 GTGAAAGTCAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACC 4502
DB 799 GTGAAAGTCAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACC 740
QY 4503 ACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAG 4562
DB 739 ACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAG 680
QY 4563 AGAAAAGTGAAGTCGAGTCGCTCGGTACC 4589
DB 679 AGAAAAGTGAAGTCGAGTCGCTCGGTACC 653
```

```
RESULT 14
US-10-102-143-8/c
; Sequence 8, Application US/10102143
```

```
; Publication No. US20030185851A1
; GENERAL INFORMATION:
; APPLICANT: Soldati, Dominique
; APPLICANT: Meisner, Markus
; TITLE OF INVENTION: TET transactivator system
; FILE REFERENCE: 04630/016001
; CURRENT APPLICATION NUMBER: US/10/102,143
; CURRENT FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 4556
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1270)..(2001)
; OTHER INFORMATION: Description of Artificial Sequence: pTetO7Sag1-GFP
US-10-102-143-8
```

```
Query Match 5.4%; Score 311; DB 16; Length 4556;
Best Local Similarity 96.9%; Pred. No. 1.5e-84;
Matches 317; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
```

```
QY 4263 CTAAACGACAGGAGGAACTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGA 4322
DB 979 CTATCGATACCGTCGACCTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGA 920
QY 4323 GTGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCC 4382
DB 919 GTGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCC 860
QY 4383 TATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAAA 4442
DB 859 TATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAGAGAAA 800
QY 4443 GTGAAAGTCAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACC 4502
DB 799 GTGAAAGTCAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACC 740
QY 4503 ACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAG 4562
DB 739 ACTCCCTATCAGTATAGAGAAAAGTGAAGTCGAGTTTACCACTCCCTATCAGTATAG 680
QY 4563 AGAAAAGTGAAGTCGAGTCGCTCGGTACC 4589
DB 679 AGAAAAGTGAAGTCGAGTCGCTCGGTACC 653
```

```
RESULT 15
US-10-102-143-6/c
; Sequence 6, Application US/10102143
; Publication No. US20030185851A1
; GENERAL INFORMATION:
; APPLICANT: Soldati, Dominique
; APPLICANT: Meisner, Markus
; TITLE OF INVENTION: TET transactivator system
; FILE REFERENCE: 04630/016001
; CURRENT APPLICATION NUMBER: US/10/102,143
; CURRENT FILING DATE: 2003-01-21
; NUMBER OF SEQ ID NOS: 31
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 6346
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1193)..(3787)
; NAME/KEY: misc feature
; LOCATION: (1102)
```

; OTHER INFORMATION: n is disclosed as an asterisk
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:
; OTHER INFORMATION: pterc07sagl-MyoA
US-10-102-143-6

Query Match 5.4%; Score 311; DB 16; Length 6346;
Best Local Similarity 96.9%; Pred. No. 1.8e-84;
Matches 317; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY	4263	CTAACGACACGAGGGAACTCGAGTTTACCACTCCCTATCATGATGATAGAGAAAAGTGAAA	4322
Db	979	CTTATCGATACCGTCGACCTCGAGTTTACCACTCCCTATCATGATGATAGAGAAAAGTGAAA	920
QY	4323	GTCGAGTTTACCACTCCCTATCATGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCC	4382
Db	919	GTCGAGTTTACCACTCCCTATCATGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCC	860
QY	4383	TATCAGTATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCATGATGATAGAGAAA	4442
Db	859	TATCAGTATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCATGATGATAGAGAAA	800
QY	4443	GTGAAAGTCGAGTTTACCACTCCCTATCATGATGATAGAGAAAAGTGAAAGTCGAGTTTACC	4502
Db	799	GTGAAAGTCGAGTTTACCACTCCCTATCATGATGATAGAGAAAAGTGAAAGTCGAGTTTACC	740
QY	4503	ACTCCCTATCATGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCATGATAG	4562
Db	739	ACTCCCTATCATGATGATAGAGAAAAGTGAAAGTCGAGTTTACCACTCCCTATCATGATAG	680
QY	4563	AGAAAAGTGAAAGTCGAGCTCGGTACC	4589
Db	679	AGAAAAGTGAAAGTCGAGCTCGGTACC	653

Search completed: August 30, 2005, 09:27:53
Job time : 2396 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 29, 2005, 19:11:58 ; Search time 609 Seconds
(without alignments)
15408.944 Million cell updates/sec

Title: US-10-613-728-1

Perfect score: 5735
Sequence: 1 ggaatcctgaaggtcacaca.....ccataagattgagtcgac 5735

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgn2_6/prodata/1/ina/5A_COMB.seq: *
2: /cgn2_6/prodata/1/ina/5B_COMB.seq: *
3: /cgn2_6/prodata/1/ina/6A_COMB.seq: *
4: /cgn2_6/prodata/1/ina/6B_COMB.seq: *
5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/1/ina/backfileseq1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1279.8	22.3	1679	2 US-08-365-486A-11	Sequence 11, Appl
2	1279.8	22.3	1679	3 US-08-880-342-11	Sequence 11, Appl
3	312.6	5.5	10728	4 US-09-376-774-5	Sequence 5, Appl
4	311.8	5.4	520	1 US-08-485-971-7	Sequence 7, Appl
5	311.8	5.4	520	1 US-08-275-876-7	Sequence 7, Appl
6	311.8	5.4	520	1 US-08-383-754-7	Sequence 7, Appl
7	311.8	5.4	520	1 US-08-485-978-7	Sequence 7, Appl
8	311.8	5.4	520	2 US-08-486-814-7	Sequence 7, Appl
9	311.8	5.4	520	2 US-08-487-472-7	Sequence 7, Appl
10	311.8	5.4	520	3 US-08-485-740-7	Sequence 7, Appl
11	311.8	5.4	520	3 US-09-162-184-7	Sequence 7, Appl
12	311.8	5.4	520	3 US-09-161-902-7	Sequence 7, Appl
13	311.8	5.4	520	3 US-09-489-777A-7	Sequence 7, Appl
14	311.8	5.4	520	5 PCT-US95-08179-7	Sequence 7, Appl
15	310.4	5.4	450	1 US-08-076-726-13	Sequence 13, Appl
16	310.4	5.4	450	1 US-08-485-971-9	Sequence 9, Appl
17	310.4	5.4	450	1 US-08-260-452-6	Sequence 6, Appl
18	310.4	5.4	450	1 US-08-275-876-9	Sequence 9, Appl
19	310.4	5.4	450	1 US-08-383-754-9	Sequence 9, Appl
20	310.4	5.4	450	1 US-08-485-978-9	Sequence 9, Appl
21	310.4	5.4	450	2 US-08-481-970-6	Sequence 6, Appl
22	310.4	5.4	450	2 US-08-486-814-9	Sequence 9, Appl
23	310.4	5.4	450	2 US-08-487-472-9	Sequence 9, Appl
24	310.4	5.4	450	2 US-08-897-719-6	Sequence 6, Appl
25	310.4	5.4	450	3 US-08-485-740-9	Sequence 9, Appl
26	310.4	5.4	450	3 US-09-162-184-9	Sequence 9, Appl
27	310.4	5.4	450	3 US-09-161-902-9	Sequence 9, Appl

c 28	310.4	5.4	450	3 US-09-163-269-6	Sequence 6, Appl
c 29	310.4	5.4	450	4 US-09-489-777A-9	Sequence 9, Appl
c 30	310.4	5.4	450	5 US-09-281-674-6	Sequence 6, Appl
c 31	310.4	5.4	450	5 PCT-US95-08179-9	Sequence 9, Appl
32	309.4	5.4	450	1 US-08-076-726-12	Sequence 12, Appl
33	309.4	5.4	450	1 US-08-485-971-8	Sequence 8, Appl
34	309.4	5.4	450	1 US-08-260-452-5	Sequence 5, Appl
35	309.4	5.4	450	1 US-08-275-876-8	Sequence 8, Appl
36	309.4	5.4	450	1 US-08-383-754-8	Sequence 8, Appl
37	309.4	5.4	450	1 US-08-485-978-8	Sequence 8, Appl
38	309.4	5.4	450	1 US-08-481-970-5	Sequence 5, Appl
39	309.4	5.4	450	2 US-08-486-814-8	Sequence 8, Appl
40	309.4	5.4	450	2 US-08-487-472-8	Sequence 8, Appl
41	309.4	5.4	450	2 US-08-897-719-5	Sequence 5, Appl
42	309.4	5.4	450	3 US-08-485-740-8	Sequence 8, Appl
43	309.4	5.4	450	3 US-09-162-184-8	Sequence 8, Appl
44	309.4	5.4	450	3 US-09-161-902-8	Sequence 8, Appl
45	309.4	5.4	450	3 US-09-163-269-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-365-486A-11

; Sequence 11, Application US/08365486A

; Patent No. 5834306

; GENERAL INFORMATION:

; APPLICANT: Webster, Keith A.

; APPLICANT: Bishopric, Nanette H.

; TITLE OF INVENTION: Tissue Specific Hypoxia Regulated

; TITLE OF INVENTION: Therapeutic Constructs

; NUMBER OF SEQUENCES: 31

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Dehlinger & Associates

; STREET: 350 Cambridge Avenue, Suite 250

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94306

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/365,486A

; FILING DATE: 23-DEC-1994

; CLASSIFICATION: 514

; ATTORNEY/AGENT INFORMATION:

; NAME: Sholtz, Charles K.

; REGISTRATION NUMBER: 38,615

; REFERENCE/DOCKET NUMBER: 8255-0018

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (415) 324-0880

; TELEFAX: (415) 324-0960

; INFORMATION FOR SEQ ID NO: 11:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1679 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; INDIVIDUAL ISOLATE: Mouse alpha MHC promoter fragment

; US-08-365-486A-11

Query Match 22.3%; Score 1279.8; DB 2; Length 1679;
Best Local Similarity 82.2%; Pred. No. 0;
Matches 1616; Conservative 0; Mismatches 57; Indels 292; Gaps 1;

QY 3767 GAATTCCTTACTATCAAAAGGAAACTAGTCGTGCACTGCAAAAGTGATGCTCTCCCT 3826
Db 1 GAATTCCTTACTATCAAAAGGAAACTAGTCGTGCACTGCAAAAGTGATGCTCTCCCT 60
QY 3827 AGACATCATGACTTTTGTCTCTGGGAGACCACTGTGGAACTTCAGAGTCTGAGAGCTA 3886
Db 61 AGACATCATGACTTTTGTCTCTGGGAGACCACTGTGGAACTTCAGAGTCTGAGAGCTA 120
QY 3887 GGAAGCTCCCTCAGCCTGGAAGCTATGAGATAGCCAGGGGTGAAAAGGAGAGAG 3946
Db 121 GGAAGCTCCCTCAGCCTGGAAGCTATGAGATAGCCAGGGGTGAAAAGGAGAGAG 180
QY 3947 CTGGAGTGGAGACTTGTGTGTGGAGCAAGGGAGAGATATTAAGCTGAAAGAGAG 4006
Db 181 CTGGAGTGGAGACTTGTGTGTGGAGCAAGGGAGAGATATTAAGCTGAAAGAGAG 240
QY 4007 TGACCTTACCCAGATTGTTCACCTTCAGATTTAAAAATTAACCTGAGGTAAAGGCT 4066
Db 241 TGACCTTACCCAGATTGTTCACCTTCAGATTTAAAAATTAACCTGAGGTAAAGGCT 300
QY 4067 GGGTAGGGGAGGTGTGTGAGACGCTCCTGTCTCTCTGATGACCTGAGGCTTTG 4126
Db 301 GGGTAGGGGAGGTGTGTGAGACGCTCCTGTCTCTCTGATGACCTGAGGCTTTG 360
QY 4127 GGGAGAGGAATGTGCCAAGACTAAAAAGGCAATGAGCCAGAGGGCGAGAGCAA 4186
Db 361 GGGAGAGGAATGTGCCAAGACTAAAAAGGCAATGAGCCAGAGGGCGAGAGCAA 420
QY 4187 CAGACCTTTTCAATGGGCAAACTTTGGGGCCGTAGTATCGATTGACAAGAACTGCGCAAT 4246
Db 421 CAGACCTTTTCAATGGGCAAACTTTGGGGCCGTAGTATCGATTGACAAGAACTGCGCAAT 480
QY 4247 CGATACCTTCTTCTTCAACGAGAGAGGAACTGAGTTTACCACCTCCCTATCAGTG 4306
Db 481 CGATACCTTCTTCTTCAACGAGAGAGGAACTGAGTTTACCACCTCCCTATCAGTG 525
QY 4307 ATGAGAGAAAGTGAAGTCGATTACCACTCCTATCAGTATGAGAGAAAGTGAAAGT 4366
Db 525 ATGAGAGAAAGTGAAGTCGATTACCACTCCTATCAGTATGAGAGAAAGTGAAAGT 525
QY 4367 CGAATTATCACTCCTATCAGTATGAGAGAAAGTGAAAGTGAAGTTTACCACTCCCTA 4426
Db 525 CGAATTATCACTCCTATCAGTATGAGAGAAAGTGAAAGTGAAGTTTACCACTCCCTA 525
QY 4427 TCAATGATAGAGAAAGTGAAGTCGATTACCACTCCTATCAGTATGAGAGAAAGT 4486
Db 525 TCAATGATAGAGAAAGTGAAGTCGATTACCACTCCTATCAGTATGAGAGAAAGT 525
QY 4487 GAAAGTCGATTACCACTCCTATCAGTATGAGAGAAAGTGAAAGTGAAGTTTACCACT 4546
Db 525 GAAAGTCGATTACCACTCCTATCAGTATGAGAGAAAGTGAAAGTGAAGTTTACCACT 525
QY 4547 TCCCTATCAGTATGAGAGAAAGTGAAAGTGAAGTCGATTACCACTCCTCAAT 4606
Db 525 TCCCTATCAGTATGAGAGAAAGTGAAAGTGAAGTCGATTACCACTCCTCAAT 548
QY 4607 TTAGGACAGCAGCATATGGATGGGATATTAAGGGGCTGAGACCTGAGAGCTGTCAGAG 4666
Db 549 TTAGGACAGCAGCATATGGATGGGATATTAAGGGGCTGAGAGCTGTCAGAG 608
QY 4667 ATTTCTCCAACTCAGATTAAGAGGAGTTCGGGTGGGGGCTTTTCACTCCACCACT 4726
Db 609 ATTTCTCCAACTCAGATTAAGAGGAGTTCGGGTGGGGGCTTTTCACTCCACCACT 668
QY 4727 CTCCCACTCAGATTAAGAGGAGTTCGGGTGGGGGCTTTTCACTCCACCACT 4786
Db 669 CTCCCACTCAGATTAAGAGGAGTTCGGGTGGGGGCTTTTCACTCCACCACT 728
QY 4787 GACAGGGGTGCTTCCACAGCCTGGAGAGTTCAGTGGCAGAGAGTTTCAACAAGAAA 4846
Db 729 GACAGGGGTGCTTCCACAGCCTGGAGAGTTCAGTGGCAGAGAGTTTCAACAAGAAA 788
QY 4847 CACTGATGCCCCCTTCCCTTAAGCTGCTTCTTCCATCTTCTCTGGGAGATGCTCTCC 4906

Db 789 CACTGATGCCCCCTTCCCTTAAGCTGCTTCTTCCATCTTCTCTGGGAGATGCTCTCC 848
QY 4907 CGTCTGATTAATCTTGGGCTCTTCCGCTCAGAGAAATTGGCCCTGCTGCTCACTGCA 4966
Db 849 CGTCTGATTAATCTTGGGCTCTTCCGCTCAGAGAAATTGGCCCTGCTGCTCACTGCA 908
QY 4967 TCTTCTCTACTGCTCCGCTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 5026
Db 909 TCTTCTCTACTGCTCCGCTGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 968
QY 5027 TTTTCACTTCACTTTTCTCCCTTCTCATTTGATTTATCATCTTCTTCTTCTTCT 5086
Db 969 TTTTCACTTCACTTTTCTCCCTTCTCATTTGATTTATCATCTTCTTCTTCTTCT 1028
QY 5087 TCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5146
Db 1029 TCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1088
QY 5147 TCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5206
Db 1089 TCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1148
QY 5207 TAAACATCTTCCAGTGGAGCAGACCTTCACTGCTGCTGCTGCTTCTTCTTCTCTCA 5266
Db 1149 TAAACATCTTCCAGTGGAGCAGACCTTCACTGCTGCTGCTGCTTCTTCTTCTCTCA 1208
QY 5267 CCCCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5326
Db 1209 CCCCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1268
QY 5327 CTACTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5386
Db 1269 CTACTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1328
QY 5387 CAGCTGTGCTCAGATTTCTTCAAGATTTCTGAAAAAGTTAACCAAGTGAAGATTTTCC 5446
Db 1329 CAGCTGTGCTCAGATTTCTTCAAGATTTCTGAAAAAGTTAACCAAGTGAAGATTTTCC 1388
QY 5447 CTGTAGCAGCAGATCAGATTTCTCCCGAAGTCAGGCTTCCAGCTCTCTTCTCTG 5506
Db 1389 CTGTAGCAGCAGATCAGATTTCTCCCGAAGTCAGGCTTCCAGCTCTCTTCTCTG 1448
QY 5507 CCAGCTCCCGGACCTTTAGCAAACTCAGGCACTTACCCACATAGACCTTGTACA 5566
Db 1449 CCAGCTCCCGGACCTTTAGCAAACTCAGGCACTTACCCACATAGACCTTGTACA 1508
QY 5567 GAGAGCAGGCACTTTACATGAGTCTGTGTGGAGAGCCATAGGCTTACGCTTAAAAAG 5626
Db 1509 GAGAGCAGGCACTTTACATGAGTCTGTGTGGAGAGCCATAGGCTTACGCTTAAAAAG 1568
QY 5627 GGCAGGGAAGTGTGTGTGTAAGAAAGTCAGAGCTTCAATAGAGCTAGCCACACAG 5686
Db 1569 GGCAGGGAAGTGTGTGTGTAAGAAAGTCAGAGCTTCAATAGAGCTAGCCACACAG 1628
QY 5687 AAATGACAGCAGATCCCTCTATCTTCCCATTAAGAGTTTGAAGT 5731
Db 1629 AAATGACAGCAGATCCCTCTATCTTCCCATTAAGAGTTTGAAGT 1673

RESULT 2
US-08-880-342-11
Sequence 11, Application US/08880342
Patent No. 6218179
GENERAL INFORMATION:
APPLICANT: Webster, Keith A.
APPLICANT: Bishopton, Nanette H.
APPLICANT: Murphy, Brian
APPLICANT: Laderoute, Keith R.
APPLICANT: Green, Christopher J.
TITLE OF INVENTION: Tissue Specific Hypoxia Regulated
NUMBER OF SEQUENCES: 37

	CORRESPONDENCE ADDRESS:	
	ADDRESSEE: Dehlinger & Associates	
	STREET: 350 Cambridge Avenue, Suite 250	
	CITY: Palo Alto	
	STATE: CA	
	COUNTRY: USA	
	ZIP: 94306	
	COMPUTER READABLE FORM:	
	MEDIUM TYPE: Floppy disk	
	COMPUTER: IBM PC compatible	
	OPERATING SYSTEM: PC-DOS/MS-DOS	
	SOFTWARE: Patent In Release #1.0, Version #1.25	
	CURRENT APPLICATION DATA:	
	APPLICATION NUMBER: US/08/880,342	
	FILING DATE: 23-JUN-1997	
	CLASSIFICATION: 514	
	PRIOR APPLICATION DATA:	
	APPLICATION NUMBER: PCT/IB95/00996	
	FILING DATE: 13-NOV-1995	
	PRIOR APPLICATION DATA:	
	APPLICATION NUMBER: US 08/365,486	
	FILING DATE: 23-DEC-1994	
	ATTORNEY/AGENT INFORMATION:	
	NAME: Sholtz, Charles K.	
	REGISTRATION NUMBER: 38,615	
	REFERENCE/DOCKET NUMBER: 8255-0018.30	
	TELECOMMUNICATION INFORMATION:	
	TELEPHONE: (415) 324-0880	
	TELEFAX: (415) 324-0960	
	INFORMATION FOR SEQ ID NO: 11:	
	SEQUENCE CHARACTERISTICS:	
	LENGTH: 1679 base pairs	
	TYPE: nucleic acid	
	STRANDEDNESS: double	
	TOPOLOGY: linear	
	MOLECULE TYPE: DNA (genomic)	
	HYPOTHETICAL: NO	
	ANTI-SENSE: NO	
	ORIGINAL SOURCE:	
	INDIVIDUAL ISOLATE: Mouse alpha MHC promoter fragment	
	US-08-880-342-11	
	Query Match	22.3%; Score 1279.8; DB 3; Length 1679;
	Best Local Similarity	82.2%; Pred. No. 0;
	Matches 1616; Conservative	0; Mismatches 57; Indels 292; Gaps 1;
Dy	3767 GAATTCCTTACATCAAGGAAAGCAGTCGTGCACCTCGCAAGTAGTGCTCCCT	3826
Dd	1 GAATTCCTTACTCATCAAGGAAAGCAGTCGTGCACCTCGCAAGTAGTGCTCCCT	60
Dy	3827 AGACATCATGACTTTGCTCTGTGGAGGCCAGCACCTGTGGAATCTCAGTCTGAGAAGTA	3886
Dd	61 AGACATCATGACTTTGCTCTGTGGAGGCCAGCACCTGTGGAATCTCAGTCTGAGAAGTA	120
Dy	3887 GGAGGCTCCCTCAAGCTGAAAGCTATGCAATAGCCAGGGTTGAAAAGGGGAAAGGAGAG	3946
Dd	121 GGAGGCTCCCTCAAGCTGAAAGCTATGCAATAGCCAGGGTTGAAAAGGGGAAAGGAGAG	180
Dy	3947 CCTGGAGATGGAGCTGTGTTGTTGGAGSCAGGGGACAATATTAAAGCTTGAAGAAGAGS	4006
Dd	181 CCTGGAGATGGAGCTGTGTTGTTGGAGSCAGGGGACAATATTAAAGCTTGAAGAAGAGS	240
Dy	4007 TGACCCTTACCAATGTTTCAACTCACCTTCAGATTAAATAAATGAGGTAAAGGCT	4066
Dd	241 TGACCCTTACCAATGTTTCAACTCACCTTCAGATTAAATAAATGAGGTAAAGGCT	300
Dy	4067 GGGTAGGGGAGTGTGTGAGACCGCTCTGTCTCTCTTCGATGCCCCTGAGGCCCTTG	4126
Dd	301 GGGTAGGGGAGTGTGTGAGACCGCTCTGTCTCTCTTCGATGCCCCTGAGGCCCTTG	360
Dy	4127 GGGAGGAGGAGATGTGCCCCAACGACTAAAAAAAGCCATGAGCCAGAGAGGGCCAGGCGAA	4186
Dd	361 GGGAGGAGGAGATGTGCCCCAACGACTAAAAAAAGCCATGAGCCAGAGAGGGCCAGGCGAA	420

QY	1487	TAGACCTTTACATGGGGCAAACTTGGGGCCGTAATATGCAATTGACAAAGACTGGCAAT	4246
Db	421	CAGACCTTTACATGGGGCAAACTTGGGGCCGTCGTCTCTGTCACTTCAGAGCCAAAG	480
QY	4247	CGATACCCCTTCTCTTTCTTAAACGACACAGAGGGAATCGAGTTTACCACTCCCTATCAAGT	4306
Db	481	GGATCAAAAGAGGAGAGCCACGACACAGAGGGAAGTGGAGGGAG-----	525
QY	4307	ATTAGAAAAAGTAAAGTCGATTTTACCACTCCTATCAAGTATGAGAAAAGTGAAAGT	4366
Db	526	-----	525
QY	4367	CGAGTTTACCACTCCCTATCAGTGATAGAAAAAGTGAAGTTTACCACTCCCTA	4426
Db	526	-----	525
QY	4427	TCAGTGATAGAAAAAGTGAAGTCGATTTTACCACTCCTATCAGTATGAGAAAAAGT	4486
Db	526	-----	525
QY	4487	GAAAGTCGATTTACCACTCCTATCAGTGATAGAAAAAGTGAAGTCCAGTTTACCAC	4546
Db	526	-----	525
QY	4547	TCCTTATCAGTGATAGAAAAAGTGAAGTCGAGTCGAGTACGACGAGGACTCCAAT	4606
Db	526	-----GGTCCACGACGAGGACTTCCAAAT	548
QY	4607	TTAGCAGACGACATATGGATGGGATATAAAGGGCTGAGACATGAGAGCTGTCAAG	4666
Db	549	TTAGCAGACGAGGACTATGGGATGGGATATAAAGGGCTGAGACATGAGAGCTGTCAAG	608
QY	4667	ATTTCCTCAACCCAGGTAAAGGGAATTTGGGATGGGGGCTCTTCAACCACACGAACT	4726
Db	609	ATTTCCTCAACCCAGGTAAAGGGAATTTGGGATGGGGGCTCTTCAACCACACGAACT	668
QY	4727	CTCCCACTTAGAAGAAATGAGCTCTTCTGGAAGTGGGGTTGACGGCGGTCAAGATCT	4786
Db	669	CTCCCACTTAGAAGAAATGAGCTCTTCTGGAAGTGGGGTTGACGGCGGTCAAGATCT	728
QY	4787	GACAGGGTGGCTTCCACACAGCCTGGGAAGTCTCAGTGGCAGGAGGTTTCCAAAGAAA	4846
Db	729	GACAGGGTGGCTTCCACACAGCCTGGGAAGTCTCAGTGGCAGGAGGTTTCCAAAGAAA	788
QY	4847	CACGTGATGCCCTTCCCTTAAGCTGTCTTCTCATTTCTCTCTGGGATGCTCTTCCC	4906
Db	789	CACGTGATGCCCTTCCCTTAAGCTGTCTTCTCATTTCTCTCTGGGATGCTCTTCCC	848
QY	4907	CGTCTGGTTTATCTTGGCTCTTCGTTCCAGCAAGATTTGCCCTGTGCTCCACTCCA	4966
Db	849	CGTCTGGTTTATCTTGGCTCTTCGTTCCAGCAAGATTTGCCCTGTGCTGCTCACTCCA	908
QY	4967	TCCTTCTCTAAGTCTCCGAGCCTTGGCTTGCCTCTTTCGAGTGTCTTCTTTCACCCA	5026
Db	909	TCCTTCTCTAAGTCTCCGAGCCTTGGCTTGCCTCTTTCGAGTGTCTTCTTTCACCCA	968
QY	5027	TTTTCTACCTCACCTTTTCCCTTCTCATTTGTATCATTCCTTCTTCTTCTTCTCT	5086
Db	969	TTTTCTACCTCACCTTTTCCCTTCTCATTTGTATCATTCCTTCTTCTTCTTCTCT	1028
QY	5087	TCCTTCCCTTCCCTTCTTCTTCTTCTTCTTCCGCTTCCCTTCTTCTTCTTCTTCT	5146
Db	1029	TCCTTCCCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1088
QY	5147	TCCTTCCCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	5206
Db	1089	TCCTTCCCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1148
QY	5207	TAAACATCTTTCAGAGGACACAGCTTTCAGTGTGAGGATCTTCTTCTTCTTCTTCT	5266
Db	1149	TAAACATCTTTCAGAGGACACAGCTTTCAGTGTGAGGATCTTCTTCTTCTTCTTCT	1208

QY 5267 CCCCCTGGCTTGTCTGTTCCATCTGTCAGAGATCTTAGATTGTCTCCAGCCTCTG 5326
DB 1209 CCCCCTGGCTTGTCTGTTCCATCTGTCAGAGATCTTAGATTGTCTCCAGCCTCTG 1268
QY 5327 CTACTCTCTTCTGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5386
DB 1269 CTACTCTCTCTCTGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1328
QY 5387 CAGCTGTGTCTCAATCTTCAAGATCTCTGAAAAGTTAAACGAGTGAGATGTTTCC 5446
DB 1329 CAGCTGTGTCTCAATCTTCAAGATCTCTGAAAAGTTAAACGAGTGAGATGTTTCC 1388
QY 5447 CTGTAGACAGAGATCAAGATCTCTCCGAAAGTCAAGCTTCAAGCTCTCTCTCTG 5506
DB 1389 CTGTAGACAGAGATCAAGATCTCTCCGAAAGTCAAGCTTCAAGCTCTCTCTCTG 1448
QY 5507 CCGCTCTCCGCGCACTTGTAGCAACCTCAAGCACTTCAAGCTTCAAGCTTCAAGCA 5566
DB 1449 CCGCTCTCCGCGCACTTGTAGCAACCTCAAGCACTTCAAGCTTCAAGCTTCAAGCA 1508
QY 5567 GAGAGAGAGCACTTGTAGCAACCTCAAGCACTTCAAGCTTCAAGCTTCAAGCA 5626
DB 1509 GAGAGAGAGCACTTGTAGCAACCTCAAGCACTTCAAGCTTCAAGCTTCAAGCA 1568
QY 5627 GCGAGGAGAGTGTGTGTAGGAAAGTCAAGCACTTCAATGAAAGCTTCAAGCA 5686
DB 1569 GCGAGGAGAGTGTGTGTAGGAAAGTCAAGCACTTCAATGAAAGCTTCAAGCA 1628
QY 5687 AAATGACAGACAGATCCCTCTATCTCCCATGAGTTGAGT 5731
DB 1629 AAATGACAGACAGATCCCTCTATCTCCCATGAGTTGAGT 1673

RESULT 3

US-09-376-774-5
Sequence 5, Application US/09376774
Patent No. 6759236
GENERAL INFORMATION:
APPLICANT: Fung, Yuen Kai
APPLICANT: Gomer, Charles
APPLICANT: T'Ang, Anne
TITLE OF INVENTION: Methods To Enhance And Confine Expression
TITLE OF INVENTION: Of Genes
FILE REFERENCE: D6087
CURRENT APPLICATION NUMBER: US/09/376,774
CURRENT FILING DATE: 2003-03-21
PRIOR APPLICATION NUMBER: 60/096,947
PRIOR FILING DATE: 1998-08-18
NUMBER OF SEQ ID NOS: 5
SEQ ID NO 5
LENGTH: 10728
TYPE: DNA
ORGANISM: Unknown
FEATURE:
NAME/KEY: misc_feature
OTHER INFORMATION: recombinant vector pDATH-TNF?

US-09-376-774-5
Query March 5.5%; Score 312.6; DB 4; Length 10728;
Best Local Similarity 97.2%; Pred. No. 7.7e-82;

Matches 318; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 4265 AACGACAGAGAGAACTCGATTACACTCCCTATCAGTATGATGAGAAAAGTGAAGT 4324
DB 212 ATCCGACAGAGAGAGTGTGCGAGTTTACCACTCCCTATCAGTATGAGAAAAGTGAAGT 271
QY 4325 CGAGTTTACCACTCCCTATCAGTATGATGAGAAAAGTGAAGTGAAGTTTACCACTCCCTA 4384
DB 272 CGAGTTTACCACTCCCTATCAGTATGATGAGAAAAGTGAAGTGAAGTTTACCACTCCCTA 331
QY 4385 TCAGTATGAGAGAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATGAGAAAAGT 4444
DB 332 TCAGTATGAGAGAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATGAGAAAAGT 391

QY 4445 GAAAGTCGAGTTTACCACTCCCTATCAGTATGAGAAAAGTGAAGTGAAGTTTACCA 4504
DB 392 GAAAGTCGAGTTTACCACTCCCTATCAGTATGAGAAAAGTGAAGTGAAGTTTACCA 451
QY 4505 TCCTTATCAGTATGAGAAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATGAG 4564
DB 452 TCCTTATCAGTATGAGAAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATGAG 511
QY 4565 AAAAGTGAAGTGAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATGAG 4591
DB 512 AAAAGTGAAGTGAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATGAG 538

RESULT 4

US-08-485-971-7
Sequence 7, Application US/08485971
Patent No. 5589362
GENERAL INFORMATION:
APPLICANT: Bujard, Hermann
APPLICANT: Gossen, Manfred
APPLICANT: Hillen, Wolfgang
APPLICANT: Heibl, Vera
APPLICANT: Schnappinger, Dirk
TITLE OF INVENTION: Tetracycline-Regulated Transcriptional
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESSES:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,971
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/383,754
FILING DATE: 03-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,452
FILING DATE: 14-JUNE-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/076,327
FILING DATE: 14-JUNE-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/076,726
FILING DATE: 14-JUNE-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Decont, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-009CP7
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid

STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-485-971-7

Query Match 5.4%; Score 311.8; DB 1; Length 520;
Best Local Similarity 99.4%; Pred. No. 1.7e-82;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4277 GGAACCTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTTTACCAAC 4336
DB 63 GGATCTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTTTACCAAC 122
QY 4337 TCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAG 4396
DB 123 TCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAG 182
QY 4397 AAAAGTGAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTT 4456
DB 183 AAAAGTGAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTT 242
QY 4457 TACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAG 4516
DB 243 TACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAG 302
QY 4517 ATAGAGAAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTT 4576
DB 303 ATAGAGAAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTT 362
QY 4577 CGAGCTCGGTACCG 4591
DB 363 CGAGCTCGGTACCG 377

RESULT 5

US-08-275-876-7
Sequence 7, Application US/08275876
Patent No. 5654168
GENERAL INFORMATION:
APPLICANT: Bujard, Hermann
APPLICANT: Gossen, Manfred
TITLE OF INVENTION: Tetracycline-Inducible Transcriptional
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHYE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/275,876
FILING DATE:
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US
FILING DATE: 01-JULY-94
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-009CP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs

TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-275-876-7

Query Match 5.4%; Score 311.8; DB 1; Length 520;
Best Local Similarity 99.4%; Pred. No. 1.7e-82;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4277 GGAACCTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTTTACCAAC 4336
DB 63 GGATCTCGAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTTTACCAAC 122
QY 4337 TCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAG 4396
DB 123 TCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAG 182
QY 4397 AAAAGTGAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTT 4456
DB 183 AAAAGTGAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTT 242
QY 4457 TACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAG 4516
DB 243 TACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAG 302
QY 4517 ATAGAGAAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTT 4576
DB 303 ATAGAGAAAAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTT 362
QY 4577 CGAGCTCGGTACCG 4591
DB 363 CGAGCTCGGTACCG 377

RESULT 6

US-08-383-754-7
Sequence 7, Application US/08383754
Patent No. 5789156
GENERAL INFORMATION:
APPLICANT: Bujard, Hermann
APPLICANT: Gossen, Manfred
TITLE OF INVENTION: Tetracycline-Regulated Transcriptional
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHYE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/383,754
FILING DATE:
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-94
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-94
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,452
FILING DATE: 14-JUNE-94
CLASSIFICATION: 436

```

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/076,327
FILING DATE: 14-JUNE-93
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/076,726
FILING DATE: 14-JUNE-93
CLASSIFICATION: 436
ATTORNEY/AGENT INFORMATION:
NAME: DeConci, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-009CP2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-383-754-7
```

```

Query Match 5.4%; Score 311.8; DB 1; Length 520;
Best Local Similarity 99.4%; Pred. No. 1.7e-82;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 4277 GGAAGTGAATTTTACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTTTACCAC 4336
DB 63 GGATCTCGAGTTTACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTTTACCAC 122
QY 4337 TCCTATCAGTATAGAGAAAGTGAATTCAGTATAGAGAAAGTGAATTCAGTTTACCAC 4396
DB 123 TCCTATCAGTATAGAGAAAGTGAATTCAGTATAGAGAAAGTGAATTCAGTTTACCAC 182
QY 4397 AAAAGTGAATTCAGTTTACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTTT 4456
DB 183 AAAAGTGAATTCAGTTTACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTTT 242
QY 4457 TACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTATAGAGAAAGTGAATTCAGTT 4516
DB 243 TACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTATAGAGAAAGTGAATTCAGTT 302
QY 4517 ATAGAGAAAGTGAATTCAGTTTACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTT 4576
DB 303 ATAGAGAAAGTGAATTCAGTTTACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTT 362
QY 4577 CGAGCTCGTACCG 4591
DB 363 CGAGCTCGTACCG 377
```

```

RESULT 7
US-08-485-978-7
Sequence 7, Application US/08485978
Patent No. 5814618
GENERAL INFORMATION:
APPLICANT: Bujard, Hermann
APPLICANT: Gosen, Manfred
TITLE OF INVENTION: Methods for Regulating Gene Expression
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
```

```

SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,978
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/383,754
FILING DATE: 03-FEB-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,452
FILING DATE: 14-JUNE-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/076,327
FILING DATE: 14-JUNE-1993
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/076,726
FILING DATE: 14-JUNE-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: DeConci, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-009CP6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-485-978-7
```

```

Query Match 5.4%; Score 311.8; DB 1; Length 520;
Best Local Similarity 99.4%; Pred. No. 1.7e-82;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 4277 GGAAGTGAATTTTACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTTTACCAC 4336
DB 63 GGATCTCGAGTTTACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTTTACCAC 122
QY 4337 TCCTATCAGTATAGAGAAAGTGAATTCAGTATAGAGAAAGTGAATTCAGTTTACCAC 4396
DB 123 TCCTATCAGTATAGAGAAAGTGAATTCAGTATAGAGAAAGTGAATTCAGTTTACCAC 182
QY 4397 AAAAGTGAATTCAGTTTACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTTT 4456
DB 183 AAAAGTGAATTCAGTTTACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTTT 242
QY 4457 TACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTATAGAGAAAGTGAATTCAGTT 4516
DB 243 TACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTATAGAGAAAGTGAATTCAGTT 302
QY 4517 ATAGAGAAAGTGAATTCAGTTTACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTT 4576
DB 303 ATAGAGAAAGTGAATTCAGTTTACCACTCCCTATCAGTATAGAGAAAGTGAATTCAGTT 362
QY 4577 CGAGCTCGTACCG 4591
DB 363 CGAGCTCGTACCG 377
```

RESULT 8
US-08-486-814-7
Sequence 7, Application US/08486814
Patent No. 5865755
GENERAL INFORMATION:
APPLICANT: Bujard, Hermann
TITLE OF INVENTION: Animals Transgenic for a Tetracycline-
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,814
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/383,754
FILING DATE: 03-FEB-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,452
FILING DATE: 14-JUNE-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/076,726
FILING DATE: 14-JUNE-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 800
ATTORNEY/AGENT INFORMATION:
NAME: Deconti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-009CP4
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ. ID NO.: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-08-486-814-7

Query Match 5.4%; Score 311.8; DB 2; Length 520;
Best Local Similarity 99.4%; Pred. No. 1.7e-82;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4277 GGAATCGAGTTTACCACTCCCTATAGATAGAGAAAAGTGAAGTTGAGTTTACCCAC 4336
DB 63 GGATTCGAGTTTACCACTCCCTATAGATAGAGAAAAGTGAAGTTGAGTTTACCCAC 122

QY 4337 TCCCTATCAGTATAGAGAAAAGTGAAGTCAGTTTACCACTCCCTATCAGTATAGAG 4396
DB 123 TCCCTATCAGTATAGAGAAAAGTGAAGTCAGTTTACCACTCCCTATCAGTATAGAG 182
QY 4397 AAAAGTGAAGTCAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCAGTT 4456
DB 183 AAAAGTGAAGTCAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCAGTT 242
QY 4457 TACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCAGTTTACCACTCCCTATCAGTG 4516
DB 243 TACCACTCCCTATCAGTATAGAGAAAAGTGAAGTCAGTTTACCACTCCCTATCAGTG 302
QY 4517 ATAGAGAAAAGTGAAGTCAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGT 4576
DB 303 ATAGAGAAAAGTGAAGTCAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGT 362
QY 4577 CGAGCTCGGTACCG 4591
DB 363 CGAGCTCGGTACCG 377

RESULT 9
US-08-487-472-7
Sequence 7, Application US/08487472
Patent No. 5912411
GENERAL INFORMATION:
APPLICANT: Bujard, Hermann
TITLE OF INVENTION: Animal Transgenic for a Tetracycline- Inducible Transcription
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,472
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/383,754
FILING DATE: 03-FEB-1995
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-94
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,452
FILING DATE: 14-JUNE-1994
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/076,327
FILING DATE: 14-JUNE-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/076,726
FILING DATE: 14-JUNE-1993
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Deconti, Giulio A. Jr.

RESULT 10
 US-08-485-740-7
 Sequence 7, Application US/08485740
 Patent No. 6004941
 GENERAL INFORMATION:
 APPLICANT: Bujard, Hermann
 APPLICANT: Gossen, Manfred
 TITLE OF INVENTION: Methods for Regulating Gene Expression
 NUMBER OF SEQUENCES: 28
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: LAHIVE & COCKFIELD
 STREET: 60 State Street, Suite 510
 CITY: Boston
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02109-1875
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/485,740
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/383,754
 FILING DATE: 03-FEB-1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:

D6	GGATTCCTGAGTTTACCACTCCCTCATCAGTAGTAGAATAAGTAAGTCAGATTACCAC	122
OY	TCCCTATCAGTAGTAGAAAAAGTGAAGTCAGATTACCACTCCCTATCAGTAGTAGAG	4396
D6	TCCCTATCAGTAGTAGAAAAAGTGAAGTCAGATTACCACTCCCTATCAGTAGTAGAG	182
OY	AAAAGTGAAAGTCAGATTACCCTCCCTATCAGTAGTAGAAAAAGTGAAGTCAGATT	4456
D6	AAAAGTGAAAGTCAGATTACCCTCCCTATCAGTAGTAGAAAAAGTGAAGTCAGATT	242
OY	TACCACTCCCTATCAGTAGTAGAAAAAGTGAAGTCAGATTACCACTCCCTATCAGTG	4516
D6	TACCACTCCCTATCAGTAGTAGAAAAAGTGAAGTCAGATTACCACTCCCTATCAGTG	302
OY	ATAGAAGAAAGTGAAGTCAGATTACCACTCCCTATCAGTAGTAGAAAAAGTGAAGT	4576
D6	ATAGAAGAAAGTGAAGTCAGATTACCACTCCCTATCAGTAGTAGAAAAAGTGAAGT	362
OY	CGAGTCGTACCAG	4591
D6	CGAGTCGTACCAG	377

```

>>> RESULT 11
>>> US-09-162-184-7
>>> Sequence 7, Application US/09162184A
>>> Patent No. 6136954
>>> GENERAL INFORMATION:
>>> APPLICANT: Bujard, Hermann
>>> Gosse, Manfred
>>> TITLE OF INVENTION: Retriacycline-Inducible Transcriptional
>>> Activator Fusion Proteins

```

NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/162,184A
FILING DATE: 28-Sep-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/383,754
FILING DATE: 03-FEB-1995
APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-1994
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-1994
APPLICATION NUMBER: US 08/260,452
FILING DATE: 14-JUNE-1994
APPLICATION NUMBER: US 08/076,327
FILING DATE: 14-JUNE-1993
APPLICATION NUMBER: US 08/076,726
FILING DATE: 14-JUNE-1993
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-0096CN
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-162-184-7
Query Match 5.4%; Score 311.8; DB 3; Length 520;
Best Local Similarity 99.4%; Pred. No. 1.7e-82;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

DB 363 CGAGCTCGGTACCCG 377
RESULT 12
US-09-161-902-7
Sequence 7, Application US/09161902
Patent No. 6242667
GENERAL INFORMATION:
APPLICANT: Bujard, Hermann
APPLICANT: Gossen, Manfred
TITLE OF INVENTION: Animal Transgenic for a Tetracycline- Inducible Transcription
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSER: LAHIVE & COCKFIELD
STREET: 60 State Street, Suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/161,902
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/487,472
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-94
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/260,452
FILING DATE: 14-JUNE-1994
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/076,327
FILING DATE: 14-JUNE-1993
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/076,726
FILING DATE: 14-JUNE-1993
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-0096P3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-5941
TELEFAX: (617)227-7400
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-161-902-7
Query Match 5.4%; Score 311.8; DB 3; Length 520;
Best Local Similarity 99.4%; Pred. No. 1.7e-82;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Db 63 GGAATCTCGAGTTTACCACTCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTTTACAC 122
QY 4337 TCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTGTACCACTCCCTATCAGTGTAGAG 4336
Db 123 TCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTGTACCACTCCCTATCAGTGTAGAG 182
QY 4397 AAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTT 4456
Db 183 AAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTT 242
QY 4457 TACCACTCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAG 4516
Db 243 TACCACTCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAG 302
QY 4517 ATAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAGAAAAGTGAAGT 4576
Db 303 ATAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAGAAAAGTGAAGT 362
QY 4577 CGAGCTCGGTACCG 4591
Db 363 CGAGCTCGGTACCG 377

RESULT 13
US-09-489-777A-7
Sequence 7, Application US/09489777A
Patent No. 6271348
GENERAL INFORMATION:
APPLICANT: Bujard, Hermann
Gossen, Manfred
TITLE OF INVENTION: Tetracycline-Inducible Transcriptional
Inhibitor Fusion Proteins
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 28 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/489,777A
FILING DATE: 24-Jan-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/162,184
FILING DATE: 28-SEP-1998
APPLICATION NUMBER: US 08/485,978
FILING DATE: 07-JUN-1995
APPLICATION NUMBER: US 08/383,754
FILING DATE: 03-FEB-1995
APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-1994
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-1994
APPLICATION NUMBER: US 08/260,452
FILING DATE: 14-JUNE-1994
APPLICATION NUMBER: US 08/076,327
FILING DATE: 14-JUNE-1993
APPLICATION NUMBER: US 08/076,726
FILING DATE: 14-JUNE-1993
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-009C6CNDV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400

TELEFAX: (617) 742-4214
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
SEQUENCE DESCRIPTION: SEQ ID NO: 7:
US-09-489-777A-7

Query Match 5.4%; Score 311.8; DB 3; Length 520;
Best Local Similarity 99.4%; Pred. No. 1,7e-82;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4277 GGAATCTCGAGTTTACCACTCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAG 4336
Db 63 GGAATCTCGAGTTTACCACTCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAG 122
QY 4337 TCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAG 4336
Db 123 TCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAG 182
QY 4397 AAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAG 4456
Db 183 AAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAG 302
QY 4457 TACCACTCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAG 4516
Db 243 TACCACTCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAG 362
QY 4517 ATAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAG 4576
Db 303 ATAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAGAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTGTAGAG 362
QY 4577 CGAGCTCGGTACCG 4591
Db 363 CGAGCTCGGTACCG 377

RESULT 14
PCT-US95-08179-7
Sequence 7, Application PC/TUS9508179
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Tetracycline-Regulated Transcriptional
Inhibitor Fusion Proteins
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/08179
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: To be assigned
FILING DATE: 07-JUN-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/383,754
FILING DATE: 03-FEB-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/275,876
FILING DATE: 15-JULY-94
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/270,637
FILING DATE: 01-JULY-94
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A. Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-009C2PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)227-5941
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 520 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA
PCT-US95-08179-7

Query Match 5.4%; Score 311.8; DB 5; Length 520;
Best Local Similarity 99.4%; Pred. No. 1.7e-82;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4277 GGAACCTCGATTACCACTCCCTATCAGTATGAGAAAAGTGAAGTTCACAC 4336
DB 63 GGATCTCGATTACCACTCCCTATCAGTATGAGAAAAGTGAAGTTCACAC 122
QY 4337 TCCTTATCAGTATGAGAAAAGTGAAGTTCACCTCCCTATCAGTATGAG 4396
DB 123 TCCCTATCAGTATGAGAAAAGTGAAGTTCACCTCCCTATCAGTATGAG 182
QY 4397 AAAAGTGAAGTTCACCACTCCCTATCAGTATGAGAAAAGTGAAGTTCAGTT 4456
DB 183 AAAAGTGAAGTTCACCACTCCCTATCAGTATGAGAAAAGTGAAGTTCAGTT 242
QY 4457 TACCACTCCCTATCAGTATGAGAAAAGTGAAGTTCACCTCCCTATCAGTTG 4516
DB 243 TACCACTCCCTATCAGTATGAGAAAAGTGAAGTTCACCTCCCTATCAGTTG 302
QY 4517 ATGAGAAAAGTGAAGTTCACCTCCCTATCAGTATGAGAAAAGTGAAGTTCAGTT 4576
DB 303 ATGAGAAAAGTGAAGTTCACCTCCCTATCAGTATGAGAAAAGTGAAGTTCAGTT 362
QY 4577 CGAGCTCGTACCG 4591
DB 363 CGAGCTCGTACCG 377

RESULT 15
US-08-076-726-13/c
Sequence 13, Application US/08076726
Patent No. 5464758
GENERAL INFORMATION:
APPLICANT: Gossen, Manfred
APPLICANT: Bujard, Hermann
TITLE OF INVENTION: Tight Control of Gene Expression in
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox
STREET: 1100 New York Avenue, N.W.
CITY: Washington
STATE: District of Columbia
COUNTRY: United States of America
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/076,726
FILING DATE: 14-JUN-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Bemdond, Robert W.
REGISTRATION NUMBER: 32,893
REFERENCE/DOCKET NUMBER: 0942.2490001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)371-2600
TELEFAX: (202)371-2540
TELEX: 248636 SSX
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 450 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: linear
US-08-076-726-13

Query Match 5.4%; Score 310.4; DB 1; Length 450;
Best Local Similarity 99.7%; Pred. No. 4.1e-82;
Matches 311; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 4280 ACTGAGTTTACCACTCCCTATCAGTATGAGAAAAGTGAAGTTCACCTCC 4339
DB 325 ACTGAGTTTACCACTCCCTATCAGTATGAGAAAAGTGAAGTTCACCTCC 266
QY 4340 CTATCAGTATGAGAAAAGTGAAGTTCACCTCCCTATCAGTATGAGAAA 4399
DB 265 CTATCAGTATGAGAAAAGTGAAGTTCACCTCCCTATCAGTATGAGAAA 206
QY 4400 AGTGAAGTTCACCACTCCCTATCAGTATGAGAAAAGTGAAGTTCAGTTAC 4459
DB 205 AGTGAAGTTCACCACTCCCTATCAGTATGAGAAAAGTGAAGTTCAGTTAC 146
QY 4460 CACTCCCTATCAGTATGAGAAAAGTGAAGTTCACCTCCCTATCAGTATG 4519
DB 145 CACTCCCTATCAGTATGAGAAAAGTGAAGTTCACCTCCCTATCAGTATG 86
QY 4520 GAGAAAAGTGAAGTTCACCTCCCTATCAGTATGAGAAAAGTGAAGTTCAG 4579
DB 85 GAGAAAAGTGAAGTTCACCTCCCTATCAGTATGAGAAAAGTGAAGTTCAG 26
QY 4580 GCTCGTACCG 4591
DB 25 GCTCGTACCG 14

Search completed: August 30, 2005, 03:35:50
Job time : 655 secs

This Page Blank (uspto)

XX The present sequence is that of MHCmtnteto, an inducible, cardiac-
CC preferred promoter sequence derived from the mouse alpha-myosin heavy
CC chain promoter binding. The native sequence was modified by insertion of
CC a 7-repeat Teco binding site sequence. The promoter comprises a responder
CC locus that is a copy number dependent, position independent locus in
CC which various transgenes can be inserted. When uninduced, these
CC transgenes are silent. When induced, the transgenes are very active.
CC These genes can then be turned off using the inducible system. The
CC promoter is useful for expressing operably linked sequences in a cardiac
CC tissue-preferred expression pattern. Expression cassettes, host cells and
CC transgenic animals are provided. The transgenic animals exhibit inducible
CC cardiac-preferred expression of a nucleotide sequence of interest, e.g.
CC Bcl2 or glycogen synthase kinase 3-beta. These animals may have an
CC altered susceptibility to cardiopathology and may be useful for
CC identifying anti-cardiopathic compounds. The cardiopathology is
CC especially a cardiomyopathy such as familial hypertrophic
CC cardiomyopathies, dilated cardiomyopathies, peripartum cardiomyopathy,
CC restrictive cardiomyopathies, ischaemic heart disease, angina pectoris,
CC myocardial infarction, hypertensive heart disease and endocarditis
CC (claimed).

XX Sequence 5735 BP; 1458 A; 1455 C; 1545 G; 1277 T; 0 U; 0 Other;

Query Match 100.0%; Score 5735; DB 12; Length 5735;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 5735; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 GGATCTCGCAAGTGCACAAAGGGTCTCCACCCACAGGTCCTTACTCAATTTCACT
DB 1 GGATCTCGCAAGTGCACAAAGGGTCTCCACCCACAGGTCCTTACTCAATTTCACT
QY 61 TTCCATGCTGTTCTCTCAATAGCTGCTCCAGAGCTATTTGGACTTTGTTTAT
DB 61 TTCCATGCTGTTCTCTCAATAGCTGCTCCAGAGCTATTTGGACTTTGTTTAT
QY 121 TTCAAAAGGCTGATAGAGATGATCTTGTCTACCCAGCTCTAAGGGTCCGCTGA
DB 121 TTCAAAAGGCTGATAGAGATGATCTTGTCTACCCAGCTCTAAGGGTCCGCTGA
QY 181 AGCCCTCAGACCTGAGCCTTTGCAAGCCCTTTAGTGGAACAAATTAAGCAATTT
DB 181 AGCCCTCAGACCTGAGCCTTTGCAAGCCCTTTAGTGGAACAAATTAAGCAATTT
QY 241 TCCTTTAAAGCCAAATCTGCTCTAGACTCTTCTTCTGACCTCGGTCCTGAGCTCT
DB 241 TCCTTTAAAGCCAAATCTGCTCTAGACTCTTCTTCTGACCTCGGTCCTGAGCTCT
QY 301 AGGGTGGGAGGTGGGCTTGAAGAAAGAGTGGGGAAGTGCAAAAGCCGATCCCTAG
DB 301 AGGGTGGGAGGTGGGCTTGAAGAAAGAGTGGGGAAGTGCAAAAGCCGATCCCTAG
QY 361 GGCCTGTGAGTTCGAGCCTTCCCTGTACAGACCTGCTCATAGATCTCTCCAGCC
DB 361 GGCCTGTGAGTTCGAGCCTTCCCTGTACAGACCTGCTCATAGATCTCTCCAGCC
QY 421 AAACATAGCAAGATGATACCTCTTGTGACTTCCACAGGCCAGTACTGTCAAGTT
DB 421 AAACATAGCAAGATGATACCTCTTGTGACTTCCACAGGCCAGTACTGTCAAGTT
QY 481 GAAACAGAGTTTAAAGAAAGCTTGAACCTCACTGAACCTTGAAGCTTCACCAAGCA
DB 481 GAAACAGAGTTTAAAGAAAGCTTGAACCTCACTGAACCTTGAAGCTTCACCAAGCA
QY 541 AGCACTAGAGTCCAGCTGATAGTATCTCTAGCTGATATATGAGAGCTGGGCAC
DB 541 AGCACTAGAGTCCAGCTGATAGTATCTCTAGCTGATATATGAGAGCTGGGCAC
QY 601 AGAAGTCTGGGGGTGAGAACTGACAGTACTTTTCAGTCGGCAAGGATATACCCCC
DB 601 AGAAGTCTGGGGGTGAGAACTGACAGTACTTTTCAGTCGGCAAGGATATACCCCC
QY 661 TCAGCAGATGTATATGTCCCTTAAATGCCATCCAGGAGGTCTCTAAAGAGACATG
```

```
DB 661 TCAGCAGATGTATATGTCCCTTAAATGCCATCCAGGAGGTCTCTAAAGAGACATG
QY 721 GATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
DB 721 GATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG
QY 781 TTCCATTTAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 781 TTCCATTTAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 841 GACACACCCCTGCTGAGTATGATGATGATGATGATGATGATGATGATGATG
DB 841 GACACACCCCTGCTGAGTATGATGATGATGATGATGATGATGATGATGATG
QY 901 TTGCAAGGCTTTGACCTGCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 901 TTGCAAGGCTTTGACCTGCTGAGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 961 GGCATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 961 GGCATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 1021 TCACCTGCGGGCTGCGGGGTAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 1021 TCACCTGCGGGCTGCGGGGTAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 1081 TAGTCCCGAGATGATCTGCAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG
DB 1081 TAGTCCCGAGATGATCTGCAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG
QY 1141 TGAAGACATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG
DB 1141 TGAAGACATGAGATGATGATGATGATGATGATGATGATGATGATGATGATG
QY 1201 ACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
DB 1201 ACACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT
QY 1261 GCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 1261 GCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 1321 CAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 1321 CAGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 1381 TGCAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 1381 TGCAATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 1441 AGATTTGAGGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 1441 AGATTTGAGGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 1501 ATACCTTCATCCCGTCTCCCAATTAAGCCCACTTTCTTCTAGATCAGACTGAG
DB 1501 ATACCTTCATCCCGTCTCCCAATTAAGCCCACTTTCTTCTAGATCAGACTGAG
QY 1561 AGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 1561 AGGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 1621 GCAACCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 1621 GCAACCTTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 1681 CTGTGACACAGAGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
DB 1681 CTGTGACACAGAGAGGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY 1741 ATGGGCGGGGGGGGATTTTGGGGGGGGGAGAGAGAGAGAGAGAGAGAGAGAG
```

Db 1741 ATGGCGGGGGGGGATTCCTGGGGGGGAGAGAGAGGTGAGAGAGAGCTGGAACAG 1800
Qy 1801 AGAATCTGGAAGGCTGGAAGACGATACATTAAGGGAAGAACCCAGGCTTACTTTGATG 1860
Db 1801 AGAATCTGGAAGGCTGGAAGACGATACATTAAGGGAAGAACCCAGGCTTACTTTGATG 1860
Qy 1861 TAAATCATGAAAAGACAGGAGAGAGAGAGCTGAGAGAGTGAAGAGACCCGGGGGCAAGA 1920
Db 1861 TAAATCATGAAAAGACAGGAGAGAGAGAGCTGAGAGAGTGAAGAGACCCGGGGGCAAGA 1920
Qy 1921 CATGGAAGCAGAGACAGCCAGGTTGAGCGCTCCGTGAATCAGCCTGTGAAAGGAGAG 1980
Db 1921 CATGGAAGCAGAGACAGCCAGGTTGAGCGCTCCGTGAATCAGCCTGTGAAAGGAGAG 1980
Qy 1981 CCTGTGTATAGACACCAAGACAGAGAGGTTAGGGTTAATGTGAGACAGGGAACGAAG 2040
Db 1981 CCTGTGTATAGACACCAAGACAGAGAGGTTAGGGTTAATGTGAGACAGGGAACGAAG 2040
Qy 2041 GTAGACACAGGAACAGACAGAGCGGGGAGCCAGGTAACAAAGAAATGGTCTTCTCAC 2100
Db 2041 GTAGACACAGGAACAGACAGAGCGGGGAGCCAGGTAACAAAGAAATGGTCTTCTCAC 2100
Qy 2101 CTGTGGCAGAGGCGTCATCTGTGTCCACATCTTAAATGTTCATCAGACTGACGGGC 2160
Db 2101 CTGTGGCAGAGGCGTCATCTGTGTCCACATCTTAAATGTTCATCAGACTGACGGGC 2160
Qy 2161 TGGCTTGGAGGAGGCTGGAAGAGTATGTGAGAGCCAGGGGAGACAAAGGGGCTTAAAGA 2220
Db 2161 TGGCTTGGAGGAGGCTGGAAGAGTATGTGAGAGCCAGGGGAGACAAAGGGGCTTAAAGA 2220
Qy 2221 AAGGAAGAAAGGGGCAAAACCAAGGCCACAGAAAGGGGAGAGCCAGAACTGATTAATC 2280
Db 2221 AAGGAAGAAAGGGGCAAAACCAAGGCCACAGAAAGGGGAGAGCCAGAACTGATTAATC 2280
Qy 2281 CTTCCTTGTGTGATCTTTCATAGAGAGCAGTGGGAACTCTGTGACCAACATCCCCATGA 2340
Db 2281 CTTCCTTGTGTGATCTTTCATAGAGAGCAGTGGGAACTCTGTGACCAACATCCCCATGA 2340
Qy 2341 GCGCCCACTACCCATACCAAGTTGGGCTGAGTGGCAATTCAGGTTCCCTGAGAGACAGAG 2400
Db 2341 GCGCCCACTACCCATACCAAGTTGGGCTGAGTGGCAATTCAGGTTCCCTGAGAGACAGAG 2400
Qy 2401 CTTGGCCTTGTCTCTTGGACCTGACCCAGCTGACCCCAATGTTCACAGTACCTTTGAAT 2460
Db 2401 CTTGGCCTTGTCTCTTGGACCTGACCCAGCTGACCCCAATGTTCACAGTACCTTTGAAT 2460
Qy 2461 GCGCTCAAGAGCTTGAGAACCAAGCAGTGAATATTAGGCTATGGGCTTAACTCTGAGCT 2520
Db 2461 GCGCTCAAGAGCTTGAGAACCAAGCAGTGAATATTAGGCTATGGGCTTAACTCTGAGCT 2520
Qy 2521 TGCACACAGAGGCTTCAAGTGAACCTCCAGGGACACAGCTGACAGAGTGGCTTTATCC 2580
Db 2521 TGCACACAGAGGCTTCAAGTGAACCTCCAGGGACACAGCTGACAGAGTGGCTTTATCC 2580
Qy 2581 CCAAGAGCAACCAATTTGGATGAGTGGCTGCAAAATGGGAATCAAGGTTGAATCAGGTC 2640
Db 2581 CCAAGAGCAACCAATTTGGATGAGTGGCTGCAAAATGGGAATCAAGGTTGAATCAGGTC 2640
Qy 2641 CTTTCAAGAAATCTGATGACAGACTAGACCCCTGAGAGAGAGGTTATGCTCTGCC 2700
Db 2641 CTTTCAAGAAATCTGATGACAGACTAGACCCCTGAGAGAGAGGTTATGCTCTGCC 2700
Qy 2701 CCACCCACCAATAGGGAGTGAATCTATCTTGGGGGCTGGCGACTTTGGGGAACACAC 2760
Db 2701 CCACCCACCAATAGGGAGTGAATCTATCTTGGGGGCTGGCGACTTTGGGGAACACAC 2760
Qy 2761 ATTAAGAGAGTGTGAGAGCCAGAAATGACCCGCTGTGTCTGCGACACCTCCAC 2820
Db 2761 ATTAAGAGAGTGTGAGAGCCAGAAATGACCCGCTGTGTCTGCGACACCTCCAC 2820
Qy 2821 TCTAGAGCTATATTGAGAGGTGACAGTATAGTGGTGGAGCTGTGACAGGAGAGTGT 2880
Db 2821 TCTAGAGCTATATTGAGAGGTGACAGTATAGTGGTGGAGCTGTGACAGGAGAGTGT 2880

Qy 2881 TCTGGGTGTAGAGGTGTAGAGGGAAGAGCAGAGCAGGGAGTGTGGCTTTGTCTCTGA 2940
Db 2881 TCTGGGTGTAGAGGTGTAGAGGGAAGAGCAGAGCAGGGAGTGTGGCTTTGTCTCTGA 2940
Qy 2941 CACAATGTCTACTTATGATTAAACAGGATGACCTGTAAAGACCAACATCTACGACTC 3000
Db 2941 CACAATGTCTACTTATGATTAAACAGGATGACCTGTAAAGACCAACATCTACGACTC 3000
Qy 3001 TGAAGAAGCAGAGCCCTGAGAGACAGGGGTGTCTGAGCCTTGGTGTGATGTG 3060
Db 3001 TGAAGAAGCAGAGCCCTGAGAGACAGGGGTGTCTGAGCCTTGGTGTGATGTG 3060
Qy 3061 CCACAAAGAGGAGCATGAGTGAATTAAGGCCACAGAGCGTTAGAGAGGCACTTG 3120
Db 3061 CCACAAAGAGGAGCATGAGTGAATTAAGGCCACAGAGCGTTAGAGAGGCACTTG 3120
Qy 3121 GGAAGGGGTCAAGTCTGAGAGCCCTTATCCATGAAATCTGAGGCTGGGGCCAACTG 3180
Db 3121 GGAAGGGGTCAAGTCTGAGAGCCCTTATCCATGAAATCTGAGGCTGGGGCCAACTG 3180
Qy 3181 TAAATCTCTGGGCTTGGCCAGGATTCAAAGCAGACCTTGATCTCTTGGCAGGCTGG 3240
Db 3181 TAAATCTCTGGGCTTGGCCAGGATTCAAAGCAGACCTTGATCTCTTGGCAGGCTGG 3240
Qy 3241 GGGGGAAGGAGCAACCCGCCCTTATCCCTTCTCCCTGAGCCCAAGATTAACACT 3300
Db 3241 GGGGGAAGGAGCAACCCGCCCTTATCCCTTCTCCCTGAGCCCAAGATTAACACT 3300
Qy 3301 CTGGCCTTCCCTTCCCATCCATCAGAGGTGAAGGTTGACAGAGGAGGTAAAA 3360
Db 3301 CTGGCCTTCCCTTCCCATCCATCAGAGGTGAAGGTTGACAGAGGAGGTAAAA 3360
Qy 3361 CCTACATGTCCAAACATCATGTGTGACAGATATATGGAATCAGTATGTGAGAGCAAGAA 3420
Db 3361 CCTACATGTCCAAACATCATGTGTGACAGATATATGGAATCAGTATGTGAGAGCAAGAA 3420
Qy 3421 GGAATCTGACAGGCTTAACTGGGTTAATGTGTAAGTCTGTGTGATGTGTCTG 3480
Db 3421 GGAATCTGACAGGCTTAACTGGGTTAATGTGTAAGTCTGTGTGATGTGTCTG 3480
Qy 3481 ACTGAAAACGGGACATGCTGTGACAGCTGTTCAATCTGTGAGGTTAACGACCTGA 3540
Db 3481 ACTGAAAACGGGACATGCTGTGACAGCTGTTCAATCTGTGAGGTTAACGACCTGA 3540
Qy 3541 GGTGTGTGTATTAATTGCCAAAGGCAAGTGGGTGAATCCCTTCAATGTTTAAAGAT 3600
Db 3541 GGTGTGTGTATTAATTGCCAAAGGCAAGTGGGTGAATCCCTTCAATGTTTAAAGAT 3600
Qy 3601 TGGATGATGCGCTGATCTCAAGACCAATGGAATAGATGATGATGATGATG 3660
Db 3601 TGGATGATGCGCTGATCTCAAGACCAATGGAATAGATGATGATGATGATGATG 3660
Qy 3661 CTAAAGCTAAGTGAAGGCTTTTGGAGGACACTGTCTAGAGATGTGGGCAACAGAGAC 3720
Db 3661 CTAAAGCTAAGTGAAGGCTTTTGGAGGACACTGTCTAGAGATGTGGGCAACAGAGAC 3720
Qy 3721 TACAGACATATCTGTACAGAGTAAAGAGAGAGAGAGGGGCTGTAGAAATCTCTTACTA 3780
Db 3721 TACAGACATATCTGTACAGAGTAAAGAGAGAGAGAGGGGCTGTAGAAATCTCTTACTA 3780
Qy 3781 TGAAGAGGAACATGAGTGTGACCTGCAAAATGGAATGATGTCCTTCAAGACATCAGACT 3840
Db 3781 TGAAGAGGAACATGAGTGTGACCTGCAAAATGGAATGATGTCCTTCAAGACATCAGACT 3840
Qy 3841 TGTCTCTGGGAGCAGACACTGTGAACCTTCAAGTCTGAGAGATGAGAGGCTCCCTCA 3900
Db 3841 TGTCTCTGGGAGCAGACACTGTGAACCTTCAAGTCTGAGAGATGAGAGGCTCCCTCA 3900
Qy 3901 GCGTGAAGCTATGACAGTATGCAAGGTTGAAGGGGGAGAGGAGGCTGGAGTGGAGC 3960
Db 3901 GCGTGAAGCTATGACAGTATGCAAGGTTGAAGGGGGAGAGGAGGCTGGAGTGGAGC 3960

OY	3961	TTGTGCTGTGGAGCGCAGGGGACAGTATTTAAAGCCGTGAAGAAGATGTAACCCCTAACCG	4020
Dd	3961	TGTGTGTGGAGCGAGGGGACAGATATTAAAGCTCGAAGAGAAAGTGTAACCTTAACCGAG	4020
OY	4021	TTGTTCACACTCACCCCTTCAGATATTAATAAATACTAGATTAAGGGCTGGGTATGGGAGAGTG	4080
Dd	4021	TTGTTCACACTCACCCCTTCAGATATTAATAAATACTAGATTAAGGGCTGGGTATGGGAGAGTG	4080
OY	4081	GTTGTGAGACGCCTCCGTGTCTTCTCTGTGATGCGCTGAGGCCCTTTGGGGAGAGAAATGT	4140
Dd	4081	GTTGTGAGACGCCTCCGTGTCTTCTCTGTGATGCGCTGAGGCCCTTTGGGGAGAGAAATGT	4140
OY	4141	GCCCCAAGACTAAAAAAAAGCCCATGAGACCAGAGGGGAGAGGGCAACAGACTTTCATNG	4200
Dd	4141	GCCCCAAGACTAAAAAAAAGCCCATGAGACCAGAGGGGAGAGGGCAACAGACTTTCATNG	4200
OY	4201	GCAAACTTTGGGGCCCGTAGTATCATTTGACAAAGACTCGCAATCGATACCCCTTCTTC	4260
Dd	4201	GCAAACTTTGGGGCCCGTAGTATCATTTGACAAAGACTCGCAATCATACCCCTTCTTC	4260
OY	4261	TTCTTAACGGAACAGAGGGGAACTCGAGTTTACCACTCCCTATACAGATATAGAAAAAGTGA	4320
Dd	4261	TTCTTAACGGAACAGAGGGGAACTCGAGTTTACCACTCCCTATACAGATATAGAAAAAGTGA	4320
OY	4321	AAGTCAGATTACCACTCCCTATCAGTATAGAAAAAGTGAAGTTCAGATTACCACTC	4380
Dd	4321	AAGTCAGATTACCACTCCCTATCAGTATAGAAAAAGTGAAGTTCAGATTACCACTC	4380
OY	4381	CCTATACAGTATAGAAAAAGTGAAGTTCAGATTACCACTCCCTATACAGATATAGAAA	4440
Dd	4381	CCTATACAGTATAGAAAAAGTGAAGTTCAGATTACCACTCCCTATACAGATATAGAAA	4440
OY	4441	AAGTGAAGTCAGATTACCACTCCCTATCAGTATAGAAAAAGTGAAGTTCAGATTAC	4500
Dd	4441	AAGTGAAGTCAGATTACCACTCCCTATCAGTATAGAAAAAGTGAAGTTCAGATTAC	4500
OY	4501	CCACTCCCTATACAGTATAGAAAAAGTGAAGTTCAGATTACCACTCCCTATACAGTAT	4560
Dd	4501	CCACTCCCTATACAGTATAGAAAAAGTGAAGTTCAGATTACCACTCCCTATACAGTAT	4560
OY	4561	AGAGAAAAGTGAAGTCGAGCTCGGTACAGACAGAGAGACTCCAAATTTAAGGACAGAGCA	4620
Dd	4561	AGAGAAAAGTGAAGTCGAGCTCGGTACAGACAGAGAGACTCCAAATTTAAGGACAGAGCA	4620
OY	4621	TATGGATAGGATATPAAAGGGGCTGAGACA CTGAGAGCTGTACAGATTTCTCAAACCA	4680
Dd	4621	TATGGATAGGATATPAAAGGGGCTGAGACA CTGAGAGCTGTACAGATTTCTCAAACCA	4680
OY	4681	GSTTAAGAGGAGATTTCCGGTGGGGGCTGTACCCACCAAGAACCTCCCACCTAGAA	4740
Dd	4681	GSTTAAGAGGAGATTTCCGGTGGGGGCTGTACCCACCAAGAACCTCCCACCTAGAA	4740
OY	4741	GGAAA CTGCTTCTGTGAGTGGGTTCA GGGCCGGT CAGAGATCTGA CAGGATGGCCCTT	4800
Dd	4741	GGAAA CTGCTTCTGTGAGTGGGTTCA GGGCCGGT CAGAGATCTGA CAGGATGGCCCTT	4800
OY	4801	CCACCAAGCTGGGAA GTTCTCAGTGGCAGAGAGTTTCCA CAAGAAACA CTGATATGCCCT	4860
Dd	4801	CCACCAAGCTGGGAA GTTCTCAGTGGCAGAGAGTTTCCA CAAGAAACA CTGATATGCCCT	4860
OY	4861	TCCCTTAAGCTCTTCTTTCATCTTCTCTGAGGATGCTCCTCCCGCTTTGGTTTTATC	4920
Dd	4861	TCCCTTAAGCTCTTCTTTCATCTTCTCTGAGGATGCTCCTCCCGCTTTGGTTTTATC	4920
OY	4921	TTGGCTCTTTCGTTCA GCAAGATTTGCC CTGTGCTGTCA CTCATCTTTCTTACTGT	4980
Dd	4921	TTGGCTCTTTCGTTCA GCAAGATTTGCC CTGTGCTGTCA CTCATCTTTCTTACTGT	4980
OY	4981	CTCCGTCGCTTCGCTTTCAGCAAAATTTGGCCCTGTGCTGTCA CTCATCTTTCTTACTGT	5040
Dd	4981	CTCCGTCGCTTCGCTTTCAGCAAAATTTGGCCCTGTGCTGTCA CTCATCTTTCTTACTGT	5040
OY	5041	TTTTTCCCTTCGATVTGATATCATCCCTTCCTTCCTTCCTTCCTTCCTTCCTTCCTTC	5100

[illegible]

XX Field LJ, Pasumatchi KBS;
XX WPI; 2001-102575/11.
XX Increasing proliferative potential of cardiomyocyte cell which is used
XX for screening activity of biological or pharmacological agent, involves
XX increasing the level of cyclin D2 activity in cardiomyocyte cell.
PS Example 1; Page 62-66; 67bp; English.
XX The present sequence is mouse alpha-cardiac myosin heavy chain (MHC)
XX promoter. This sequence is used in the preparation of a MHC-CYCD2 fusion
XX gene. The cyclin D2 (CYCD2) cDNA sequence is used to increase the
XX proliferative potential of a cardiomyocyte cell by increasing the level
XX of cyclin D2 activity in the cardiomyocyte cell. The transgenic animals
XX expressing cyclin D2 have sustained atrial and ventricular cardiomyocyte
XX DNA synthesis. Cardiomyocyte cells with enhanced proliferative potential
XX are useful for screening the activity of biological or pharmacological
XX agents on cardiomyocyte cells. Genetically modified cardiomyocyte cells
XX are useful for delivering therapeutics to mammals. The cells are also
XX useful to target an improvement of the contractile function of the heart
XX of the patient, for e.g. in the treatment of contractile losses due to
XX infarcts or cardiomyopathies
SQ Sequence 5443 BP; 1352 A; 1399 C; 1492 G; 1200 T; 0 U; 0 Other;
Query Match 88.0%; Score 5045; DB 5; Length 5443;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 5383; Conservative 0; Mismatches 60; Indels 292; Gaps 1;
QY 1 GGATCCGCAAGGACACAAAGGCTCCACCAACGAGGCTGCTCATATTTCACT 60
DB 1 GGATCTCGAAGTCAACAAAGGCTCTCCACCAACGAGGCTGCTCATATTTCACT 60
QY 61 TTCATGCTTGTCTCAACATGCTGCTCCACAGACTAATTTGGACTTTTAT 120
DB 61 TTCATGCTTGTCTCAACATGCTGCTCCACAGACTAATTTGGACTTTTAT 120
QY 121 TTCAAAAGGCTGATAGAGAGATGATCTTTGTCTACCAAGCTTAAGGTCCTGTA 180
DB 121 TTCAAAAGGCTGATAGAGAGATGATCTTTGTCTACCAAGCTTAAGGTCCTGTA 180
QY 181 AGGCTCAGACCTGAGACCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 181 AGGCTCAGACCTGAGACCTTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 TCCTTAAGCAAAATCTGCTCTAGACTCTTCTCTGACCTCGGTCCTGAGCTCT 300
DB 241 TCCTTAAGCAAAATCTGCTCTAGACTCTTCTCTGACCTCGGTCCTGAGCTCT 300
QY 301 AGGCTGAG 360
DB 301 AGGCTGAG 360
QY 361 GGGCTGTGAAGTTCGAGGCTTCCCTGTACAGAGAGAGAGAGAGAGAGAGAGAGAG 420
DB 361 GGGCTGTGAAGTTCGAGGCTTCCCTGTACAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 421 AAACATAGCAAGAGATGATACCTCTTGTGACTTCCCAAGGCTCAGTACCTGAGGTT 480
DB 421 AAACATAGCAAGAGATGATACCTCTTGTGACTTCCCAAGGCTCAGTACCTGAGGTT 480
QY 481 GAAACAGAGATTTAG 540
DB 481 GAAACAGAGATTTAG 540
QY 541 AGCAGCTAGAGTGCATCTGATAGTATCTTACGCTGATATATATGAGAGCTGGGAC 600
DB 541 AGCAGCTAGAGTGCATCTGATAGTATCTTACGCTGATATATATGAGAGCTGGGAC 600
QY 601 AGAAGTCTGGGAGTGAAG 660
DB 601 AGAAGTCTGGGAGTGAAG 660

DB 601 AGAAGTCTGGGAGTGAAG 660
QY 661 TCAGCAGATGATGATATGCTCCCTAGATCCATCCAGGAGAGTCTCTAAGAGACATG 720
DB 661 TCAGCAGATGATGATATGCTCCCTAGATCCATCCAGGAGAGTCTCTAAGAGACATG 720
QY 721 GGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
DB 721 GGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
QY 781 TTCACCTTAGCCAG 840
DB 781 TTCACCTTAGCCAG 840
QY 841 GACACACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
DB 841 GACACACCCCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 900
QY 901 TTGCAAGGCTTTGACCTCTGAGCTGCTGAGAGATGATGATGATGATGATGATGATGAT 960
DB 901 TTGCAAGGCTTTGACCTCTGAGCTGCTGAGAGATGATGATGATGATGATGATGATGAT 960
QY 961 GCCATCTCAAG 1020
DB 961 GCCATCTCAAG 1020
QY 1021 TCACCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
DB 1021 TCACCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1081 TAGTCCCGAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1081 TAGTCCCGAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1141 TGAAGACATGAGATGAG 1200
DB 1141 TGAAGACATGAGATGAG 1200
QY 1201 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
DB 1201 AACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 1261 GCAACAGAGATGAG 1320
DB 1261 GCAACAGAGATGAG 1320
QY 1321 CAGGAG 1380
DB 1321 CAGGAG 1380
QY 1381 TGCAGTCAAGAGATGAG 1440
DB 1381 TGCAGTCAAGAGATGAG 1440
QY 1441 AGATCTGAGAGATGAG 1500
DB 1441 AGATCTGAGAGATGAG 1500
QY 1501 ATACCTCATCCGCTCTCCCAATTAAGCCACTCTTCTCTAGATCAGACTGAGCTGC 1560
DB 1501 ATACCTCATCCGCTCTCCCAATTAAGCCACTCTTCTCTAGATCAGACTGAGCTGC 1560
QY 1561 AGCAG 1620
DB 1561 AGCAG 1620
QY 1621 GCAACCTTAAG 1680
DB 1621 GCAACCTTAAG 1680
QY 1681 CTGTGACACAG 1740
DB 1681 CTGTGACACAG 1740

Qy 1741 ATGGGCGGGGGGGGATTTCTGGGGGGGGGAGAGAGAGGTGAGAAAGGAGCTTGAAACAG 1800
Db 1741 ATGGGCGGGGGGGGAGTTCTGGGGGGGGGAGAGAGAGGTGAGAAAGGAGCTTGAAACAG 1800
Qy 1801 AGAATCTGGAAAGCGCTGGAAAACGATACCATTAAGGGAAGAACCCAGGCTTCTTTAGATG 1860
Db 1801 AGAATCTGGAAAGCGCTGGAAAACGATACCATTAAGGGAAGAACCCAGGCTTCTTTAGATG 1860
Qy 1861 TAAATCATGAAAGCAGGAGGAGGAGGAGGCTGGAGAGGTAGAGAGAGCCCGGGGGCAAGA 1920
Db 1861 TAAATCATGAAAGCAGGAGGAGGAGGAGGCTGGAGAGGTAGAGAGAGCCCGGGGGCAAGA 1920
Qy 1921 CATGGAAGCAAGGACAAAGCCAGGTTGAGCGCTCCGTGAAATCAAGCTGTGAAAGGCAAG 1980
Db 1921 CATGGAAGCAAGGACAAAGCCAGGTTGAGCGCTCCGTGAAATCAAGCTGTGAAAGGCAAG 1980
Qy 1981 CCTGTGTATGACCCAGAAACAGAGAGGCTTAGGGTTAATGTGAGACAGAGAAACAGAG 2040
Db 1981 CCTGTGTATGACCCAGAAACAGAGAGGCTTAGGGTTAATGTGAGACAGAGAAACAGAG 2040
Qy 2041 GTAGACACAGGAAACAGACAGAGAGCGGGGAGCCAGTTAACAAAGGAATGTCCTTCAC 2100
Db 2041 GTAGACACAGGAAACAGACAGAGAGCGGGGAGCCAGTTAACAAAGGAATGTCCTTCAC 2100
Qy 2101 CTGTGGCAGAGCGTCCATCTGTGTCCACATACTCTAGAAATGTTTCATCAGACTGCAAGGGC 2160
Db 2101 CTGTGGCAGAGCGTCCATCTGTGTCCACATACTCTAGAAATGTTTCATCAGACTGCAAGGGC 2160
Qy 2161 TGGCTTGGGAGGAGCTGGAAAAGATATGTGAGAGCCAGGGGAGACAAAGGGGGCTTAGGA 2220
Db 2161 TGGCTTGGGAGGAGCTGGAAAAGATATGTGAGAGCCAGGGGAGACAAAGGGGGCTTAGGA 2220
Qy 2221 AAGGAAGAGGGGCAAAACAGGCGCACACAAAGAGGGAGCCAGAACTGAGTTAATCTC 2280
Db 2221 AAGGAAGAGGGGCAAAACAGGCGCACACAAAGAGGGAGCCAGAACTGAGTTAATCTC 2280
Qy 2281 CTTCCTTGTTCATCTTCCATAGAGGAGAGTGGGAACTGTGTGACACATCCCCATGA 2340
Db 2281 CTTCCTTGTTCATCTTCCATAGAGGAGAGTGGGAACTGTGTGACACATCCCCATGA 2340
Qy 2341 GCGCCCACTACCCATACCAAGTTGGGCTGAGTGGCACTTAGGTTCCCTGAGAGACAG 2400
Db 2341 GCGCCCACTACCCATACCAAGTTGGGCTGAGTGGCACTTAGGTTCCCTGAGAGACAG 2400
Qy 2401 CCGGCGCTTGTCTCTTGACCTGACCCCAAGCTGACCCCAATGTTCTGAGTACCTTTGAT 2460
Db 2401 CCGGCGCTTGTCTCTTGACCTGACCCCAAGCTGACCCCAATGTTCTGAGTACCTTTGAT 2460
Qy 2461 GCGCTCAAGAGCTTGAGAACAGGACAGTGAATTAAGGCGATGGGCTAAACCTGAGCT 2520
Db 2461 GCGCTCAAGAGCTTGAGAACAGGACAGTGAATTAAGGCGATGGGCTAAACCTGAGCT 2520
Qy 2521 TGCACACAGAGGCTTCAAGTGAACCTCAAGGAGACAGCTGACAGAGTGGCTTTATCC 2580
Db 2521 TGCACACAGAGGCTTCAAGTGAACCTCAAGGAGACAGCTGACAGAGTGGCTTTATCC 2580
Qy 2581 CCAAAAGAGCAACATTTGGGCTAGAGTGGCTCAAAATGGGAAATGCAAGGTTGAATCAAGTC 2640
Db 2581 CCAAAAGAGCAACATTTGGGCTAGAGTGGCTCAAAATGGGAAATGCAAGGTTGAATCAAGTC 2640
Qy 2641 CTTTCAAGAAATCTGCAATGCAAGACCTTAAGACCCCTGGAGAGAGGGGTATGCTTCC 2700
Db 2641 CTTTCAAGAAATCTGCAATGCAAGACCTTAAGACCCCTGGAGAGAGGGGTATGCTTCC 2700
Qy 2701 CCACCCACATTAAGGGAGTGAATCTATCTTAGGGGCTGGGACCTTGGGGAGACACAC 2760
Db 2701 CCACCCACATTAAGGGAGTGAATCTATCTTAGGGGCTGGGACCTTGGGGAGACACAC 2760
Qy 2761 ATTACTGAGAGTGTGAGAGCCAGAAAAATGACCGCGCTGTGTCTGCTCCACCTTCCAC 2820
Db 2761 ATTACTGAGAGTGTGAGAGCCAGAAAAATGACCGCGCTGTGTCTGCTCCACCTTCCAC 2820

Qy 2821 TCTAGAGCTAATATGAGAGTGAACAGTAGATAGGGTGGAGGCTGTAGACAGGAGAGT 2880
Db 2821 TCTAGAGCTAATATGAGAGTGAACAGTAGATAGGGTGGAGGCTGTAGACAGGAGAGT 2880
Qy 2881 TCTGGGTGTGAGGGTGTAGGGGAAAGCCAGAGCAGGGGAGTCTGGCTTTGTCTCTGAA 2940
Db 2881 TCTGGGTGTGAGGGTGTAGGGGAAAGCCAGAGCAGGGGAGTCTGGCTTTGTCTCTGAA 2940
Qy 2941 CACAATGTCTACTAGTATTAACAGGACATGACCTGTAAAGACCCCAATCTTACGACTC 3000
Db 2941 CACAATGTCTACTAGTATTAACAGGACATGACCTGTAAAGACCCCAATCTTACGACTC 3000
Qy 3001 TGAAGGACAGACGCCCTGAGAGCAGGGGTGTCTTGAAGCTTGGGTCTTGATGTG 3060
Db 3001 TGAAGGACAGACGCCCTGAGAGCAGGGGTGTCTTGAAGCTTGGGTCTTGATGTG 3060
Qy 3061 CCAAAAGAGGAGCAGTGAAGTGAAGTGAAGGCCCCAGAGGCTTGAAGAGGCACTTG 3120
Db 3061 CCAAAAGAGGAGCAGTGAAGTGAAGTGAAGGCCCCAGAGGCTTGAAGAGGCACTTG 3120
Qy 3121 GGAAGGGTCAAGTCTGAGAGCCCTATCCATGGAATCTGAGCTGGGGCCAACTGTG 3180
Db 3121 GGAAGGGTCAAGTCTGAGAGCCCTATCCATGGAATCTGAGCTGGGGCCAACTGTG 3180
Qy 3181 TAAATCTGTGGCTTGCAGAGCAATGAAAGACAGCCTGATCTTGGCAGCTGGGGA 3240
Db 3181 TAAATCTGTGGCTTGCAGAGCAATGAAAGACAGCCTGATCTTGGCAGCTGGGGA 3240
Qy 3241 GGGGAAAGGAGCAACCCCCACTTATACCTTTCTCCCTCAGCCCCAGATTAAACCT 3300
Db 3241 GGGGAAAGGAGCAACCCCCACTTATACCTTTCTCCCTCAGCCCCAGATTAAACCT 3300
Qy 3301 CTGGCTTCCCTCCCTCCACTCCATCAGAGTGAAGGTTGACAGGAGGGTAAAAA 3360
Db 3301 CTGGCTTCCCTCCCTCCACTCCATCAGAGTGAAGGTTGACAGGAGGGTAAAAA 3360
Qy 3361 CCTACATGTCCAAACATCATGTGACAGATATATGAGTACATATGTGTAGAGGCAAGAA 3420
Db 3361 CCTACATGTCCAAACATCATGTGACAGATATATGAGTACATATGTGTAGAGGCAAGAA 3420
Qy 3421 GGAATCTGAGGCTTAACTGGGTTAATGTGTAATCTGTGTGCAATGTGTGTCTG 3480
Db 3421 GGAATCTGAGGCTTAACTGGGTTAATGTGTAATCTGTGTGCAATGTGTGTCTG 3480
Qy 3481 ACTGAAAGGGGATAGGCTGTGACGCTGTTCAGTTCTGTGCGTAGGTTAACAGACTGCA 3540
Db 3481 ACTGAAAGGGGATAGGCTGTGACGCTGTTCAGTTCTGTGCGTAGGTTAACAGACTGCA 3540
Qy 3541 GGTGTGTGTAAATGGCCCAAGGCAAGTGAGTATCCCTTCAATGGTTTAAAGAT 3600
Db 3541 GGTGTGTGTAAATGGCCCAAGGCAAGTGAGTATCCCTTCAATGGTTTAAAGAT 3600
Qy 3601 TGGATATGGCTGCACTTCAAGAGACATGGAATAATGGAATGCACTATATGTGTCT 3660
Db 3601 TGGATATGGCTGCACTTCAAGAGACATGGAATAATGGAATGCACTATATGTGTCT 3660
Qy 3661 CTAAAGCTAAGGTAGCAAGGCTTTGGAGGACACTGTCTAGAGATGTGGCAACAGAGAC 3720
Db 3661 CTAAAGCTAAGGTAGCAAGGCTTTGGAGGACACTGTCTAGAGATGTGGCAACAGAGAC 3720
Qy 3721 TACAGACAGTATCTTGAACAGATTAAGAGAGAGAGGGGGTGTGAATTTCTTTACTA 3780
Db 3721 TACAGACAGTATCTTGAACAGATTAAGAGAGAGAGGGGGTGTGAATTTCTTTACTA 3780
Qy 3781 TCAAAAGGAAACAGATTCGTGCACTGCAAAAGTGAATGCTCTCCCTAAGACATGACTT 3840
Db 3781 TCAAAAGGAAACAGATTCGTGCACTGCAAAAGTGAATGCTCTCCCTAAGACATGACTT 3840
Qy 3841 TGTCTCTGGGAGACAGCACTGTGGAATTCAGAGTCTGAGAGATGAGAGGCTCCCTCA 3900
Db 3841 TGTCTCTGGGAGACAGCACTGTGGAATTCAGAGTCTGAGAGATGAGAGGCTCCCTCA 3900
Qy 3901 GCCTGAAGCTATGACAGATAGCCAGGCTTGAAAGGGGAAAGGAGAGCCTGGAGTGGAGC 3960

```

Db 3901 GCGTGAAGCATATGCAAGTATGCAAGGGTGAAGAGGGGAGAGAGCTGGATGGGAGC 3960
Qy 3961 TTGTGTTGGAGGACAGGGGACAGATATTAAGCTTGAGAGAAAGGTGACCTTACCCAG 4020
Db 3961 TTGTGTTGGAGGACAGGGGACAGATATTAAGCTTGAGAGAAAGGTGACCTTACCCAG 4020
Qy 4021 TTGTGAACCTGACCTTCAGATTAATAAATTAAGAGTAAAGGCTGGGTAAGGAGGAGTG 4080
Db 4021 TTGTGAACCTGACCTTCAGATTAATAAATTAAGAGTAAAGGCTGGGTAAGGAGGAGTG 4080
Qy 4081 GTGTGAGACGCTGCTGCTCTCTCTGATGATGAGGCTTGAGGCTTTGGGAGAGAGATGT 4140
Db 4081 GTGTGAGACGCTGCTGCTCTCTCTGATGATGAGGCTTGAGGCTTTGGGAGAGAGATGT 4140
Qy 4141 GCCCAAGAGTAAAAAAGGCGATGAGAGCGAGGCGAGGCGAACAAGCTTTTCATGG 4200
Db 4141 GCCCAAGAGTAAAAAAGGCGATGAGAGCGAGGCGAGGCGAACAAGCTTTTCATGG 4200
Qy 4201 GCAACCTTGGGGGCGGTGATGATGATGACAAGAACTGCGCAATGATACCTTCTTC 4260
Db 4201 GCAACCTTGGGGGCGGTGATGATGATGACAAGAACTGCGCAATGATACCTTCTTC 4260
Qy 4261 TTCTTAAAGGAGAGAGGAACTGAGATTTCACCTCCCTATCAGTATAGAGAAAGTGA 4320
Db 4261 GGAGCAGAGACAGAGAGGAGAGTGGAGGAG----- 4291
Qy 4321 AAGTCGATTTACCATCTCCCTATCAGTATAGAGAAAGTGAAGTTTACCATCTC 4380
Db 4292 ----- 4291
Qy 4381 CCTATCAGTATAGAGAAAGTGAAGTGAAGTTTACCATCTCCCTATCAGTATAGAGAA 4440
Db 4292 ----- 4291
Qy 4441 AAGTGAAGTCGAGATTTCACATCTCCCTATCAGTATAGAGAAAGTGAAGTTTAA 4500
Db 4292 ----- 4291
Qy 4501 CCACTCCCTATCAGTATAGAGAAAGTGAAGTGAAGTTTACCATCTCCCTATCAGTAT 4560
Db 4292 ----- 4291
Qy 4561 AAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4620
Db 4292 ----- 4291
Qy 4621 TATGAGATGAGATATTAAGAGGAGCTGAGAGCTGAGAGATTTCTCAACCA 4680
Db 4329 TATGAGATGAGATATTAAGAGGAGCTGAGAGCTGAGAGATTTCTCAACCA 4388
Qy 4681 GGTAAAGGAGATTTGGGAGTGGGAGCTTTCACCAACCAAGCTTCCCACTAGAA 4740
Db 4389 GGTAAAGGAGATTTGGGAGTGGGAGCTTTCACCAACCAAGCTTCCCACTAGAA 4448
Qy 4741 GGAACCTGCTTCTGAGAGTGGGAGTGGGAGCTTTCACCAACCAAGCTTCCCACTAG 4800
Db 4449 GGAACCTGCTTCTGAGAGTGGGAGTGGGAGCTTTCACCAACCAAGCTTCCCACTAG 4508
Qy 4801 CCAACGAGCTGAGAGATTTCTCAGTGGAGAGGATTTCAACAACCAAGCTTCCCACTAG 4860
Db 4509 CCAACGAGCTGAGAGATTTCTCAGTGGAGAGGATTTCAACAACCAAGCTTCCCACTAG 4568
Qy 4861 TCCCTTAAGCTGCTTCTCAGATCTTCTCTGAGAGATGCTCTCCCGTCTTGGTTATC 4920
Db 4569 TCCCTTAAGCTGCTTCTCAGATCTTCTCTGAGAGATGCTCTCCCGTCTTGGTTATC 4628
Qy 4921 TTGGCTCTTGTCTTCAAGAGATTTGGGAGTGGGAGCTTTCACCAACCAAGCTTCCCACTAG 4980
Db 4629 TTGGCTCTTGTCTTCAAGAGATTTGGGAGTGGGAGCTTTCACCAACCAAGCTTCCCACTAG 4688
Qy 4981 CTCGAGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 5040

```

```

Db 4689 CTCGAGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 4748
Qy 5041 TTTTCTCCCTTCTGATTTGATTTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5100
Db 4749 TTTTCTCCCTTCTGATTTGATTTGATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4808
Qy 5101 CTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 5160
Db 4809 CTTCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4868
Qy 5161 CTGTGTCAGAGTGTGAGATGACACTGGGGTTCACCTTATGTAACAATCTTCCA 5220
Db 4869 CTGTGTCAGAGTGTGAGATGACACTGGGGTTCACCTTATGTAACAATCTTCCA 4928
Qy 5221 GTGAGCAGAGCTTCAAGTCTGCTGGGCTCTTCTTACCTTCTCAACCCCTGGCTTGTG 5280
Db 4929 GTGAGCAGAGCTTCAAGTCTGCTGGGCTCTTCTTACCTTCTCAACCCCTGGCTTGTG 4988
Qy 5281 CTGTTCATCTCCGATCAGAGTCTAGATGATGATGATGATGATGATGATGATGATGATGAT 5340
Db 4989 CTGTTCATCTCCGATCAGAGTCTAGATGATGATGATGATGATGATGATGATGATGATGAT 5048
Qy 5341 GCTGTCTCTCTCTGTGTCAGAGTGTGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 5400
Db 5049 GCTGTCTCTCTCTGTGTCAGAGTGTGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 5108
Qy 5401 ATTCTTCAAGATTTCTGAAAAAGTTAACAGAGTGAAGTGTTCCTGTGTAGACAGCAGA 5460
Db 5109 ATTCTTCAAGATTTCTGAAAAAGTTAACAGAGTGAAGTGTTCCTGTGTAGACAGCAGA 5168
Qy 5461 TCACGATTTCTCCGAGAGTCAAGCTTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5520
Db 5169 TCACGATTTCTCCGAGAGTCAAGCTTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5228
Qy 5521 CTCTTGAACAACCTGAGGACCTTACCCCATAGACCTTGAAGAGAGAGAGAGAGAGAGAG 5580
Db 5229 CTCTTGAACAACCTGAGGACCTTACCCCATAGACCTTGAAGAGAGAGAGAGAGAGAGAG 5288
Qy 5581 TTACATGAGTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5640
Db 5289 TTACATGAGTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5348
Qy 5641 GGTAAAGGAGATTTGGGAGTGGGAGCTTTCACCAACCAAGCTTCCCACTAGAA 5700
Db 5349 GGTAAAGGAGATTTGGGAGTGGGAGCTTTCACCAACCAAGCTTCCCACTAGAA 5408
Qy 5701 TCCCTCTATCTCCCATTAAGATTTGAGTGCAC 5735
Db 5409 TCCCTCTATCTCCCATTAAGATTTGAGTGCAC 5443

```

RESULT 3
 ABL49724
 ID ABL49724 standard; DNA; 5443 BP.
 XX
 AC ABL49724;
 XX
 DT 29-MAY-2002 (first entry)
 XX
 DE Mouse promoter DNA sequence SEQ ID NO:2.
 XX
 XX
 KM Human; p300; p300 transgenic animal; promoter; heart muscle cell;
 XX
 OS Mus musculus.
 XX
 PN M0200205633-A1.
 XX
 PD 24-JAN-2002.
 XX
 PF 13-JUL-2001; 2001WO-JP006086.
 XX
 PR 14-JUL-2000; 2000JP-00215143.

Db 1681 CTGTGACACAGAGGTCGAGGACTGTGTGTAGAGCCTCAAGATGAATGATCTCAGA 1740
QY 1741 ATGGGGGGGGGGGGGATTTCTGGGGGGGGGAGAGAGAGTGAGAGAGCCTCGAAGCAG 1800
Db 1741 ATGGGGGGGGGGGGGATTTCTGGGGGGGGGAGAGAGAGTGAGAGAGCCTCGAAGCAG 1800
QY 1801 AGAATCTGGAAGGCTGGAACGATACATTAAGGGAAGAACCCAGGCTTACTTTAGATG 1860
Db 1801 AGAATCTGGAAGGCTGGAACGATACATTAAGGGAAGAACCCAGGCTTACTTTAGATG 1860
QY 1861 TAAATCATGAAGACAGAGGGAAGGGAAGCTGAGAGAGTAGAAGAACCCCGGGGCAAGA 1920
Db 1861 TAAATCATGAAGACAGAGGGAAGGGAAGCTGAGAGAGTAGAAGAACCCCGGGGCAAGA 1920
QY 1921 CAGGAAGCAAGGACAGAGCAGGTTGAGCGCTCCGTGAATTCAGCCTGCTGAAGGCAAG 1980
Db 1921 CAGGAAGCAAGGACAGAGCAGGTTGAGCGCTCCGTGAATTCAGCCTGCTGAAGGCAAG 1980
QY 1981 CCTGTGATGAGCAGACAGAACAGCAGAGCTAGGGTTAATGTGAGACAGGGAACAGAA 2040
Db 1981 CCTGTGATGAGCAGACAGAACAGCAGAGCTAGGGTTAATGTGAGACAGGGAACAGAA 2040
QY 2041 GTAGACACAGGAACAGACAGAGACGAGGAGCCAGGTTAACAAAGAAATGCTCTTCAC 2100
Db 2041 GTAGACACAGGAACAGACAGAGACGAGGAGCCAGGTTAACAAAGAAATGCTCTTCAC 2100
QY 2101 CTGTGACAGAGGCTCATCTGTGTCCACATCTTATGAATGTTTATCAGACTGCAAGG 2160
Db 2101 CTGTGACAGAGGCTCATCTGTGTCCACATCTTATGAATGTTTATCAGACTGCAAGG 2160
QY 2161 TGGCTTGGAGGAGCTGGAAGAGTAGTAGAGCCAGGAGCAAGAGGAGGCTTAGA 2220
Db 2161 TGGCTTGGAGGAGCTGGAAGAGTAGTAGAGCCAGGAGCAAGAGGAGGCTTAGA 2220
QY 2221 AAGGAAGAGAGGAGCAAGCCAGGCAACAGAGGAGGAGCCAGAGCTGAGTTAACTC 2280
Db 2221 AAGGAAGAGAGGAGCAAGCCAGGCAACAGAGGAGGAGCCAGAGCTGAGTTAACTC 2280
QY 2281 AAGGAAGAGAGGAGCAAGCCAGGCAACAGAGGAGGAGCCAGAGCTGAGTTAACTC 2280
Db 2281 AAGGAAGAGAGGAGCAAGCCAGGCAACAGAGGAGGAGCCAGAGCTGAGTTAACTC 2280
QY 2281 CTTCCTTGTGTGATCTTTCCATAGAGGAGAGTGGGAATCTGTGTACCACTCCCATGA 2340
Db 2281 CTTCCTTGTGTGATCTTTCCATAGAGGAGAGTGGGAATCTGTGTACCACTCCCATGA 2340
QY 2341 GGGCCCACTACCATCAAGTTTGGCTGAGTGGCACTTCTAGGTTCCCTGAGAGCAG 2400
Db 2341 GGGCCCACTACCATCAAGTTTGGCTGAGTGGCACTTCTAGGTTCCCTGAGAGCAG 2400
QY 2401 CCTGGCTTGTGTCTTGTGAAGCTGACCAAGCTGACCAATGTTCTCAGTACTTTGAT 2460
Db 2401 CCTGGCTTGTGTCTTGTGAAGCTGACCAAGCTGACCAATGTTCTCAGTACTTTGAT 2460
QY 2461 GGGCTTCAAGGCTTGAAGACCAAGGAGTGAATTTAGGCAATGGGCTAACCTTGAAGCT 2520
Db 2461 GGGCTTCAAGGCTTGAAGACCAAGGAGTGAATTTAGGCAATGGGCTAACCTTGAAGCT 2520
QY 2521 TGGACACAGAGGCTCAAGTGAAGCTCAGAGGAGACAGCTGACAGAGTGGCTTATCC 2580
Db 2521 TGGACACAGAGGCTCAAGTGAAGCTCAGAGGAGACAGCTGACAGAGTGGCTTATCC 2580
QY 2581 CCAAGAGCAACATTTGGCATAGTGTGCTGAATAGGGAATCAAGTTGAATCAAGTCTC 2640
Db 2581 CCAAGAGCAACATTTGGCATAGTGTGCTGAATAGGGAATCAAGTTGAATCAAGTCTC 2640
QY 2641 CCTTCAAGATTAATCTGCACTGCAAGCTTAAGACCCCTGAGAGAGGGGTATGCTTCC 2700
Db 2641 CCTTCAAGATTAATCTGCACTGCAAGCTTAAGACCCCTGAGAGAGGGGTATGCTTCC 2700
QY 2701 CCAAGCAAGATTAAGGAGGAGTGAATCTAGAGGAGGCTGAGCACTTGGGGAAGACAC 2760
Db 2701 CCAAGCAAGATTAAGGAGGAGTGAATCTAGAGGAGGCTGAGCACTTGGGGAAGACAC 2760
QY 2761 ATTAAGTGAAGTGTGAGCCAGAAAACTGACCGGCTGTGTCTTCCCACTTCCACAC 2820
Db 2761 ATTAAGTGAAGTGTGAGCCAGAAAACTGACCGGCTGTGTCTTCCCACTTCCACAC 2820

QY 2821 TCTAGAGCTATATTTGAGAGGTGACAGTAGATAGGGTGGAGGCTGTGAGCAGGAGAGTGT 2880
Db 2821 TCTAGAGCTATATTTGAGAGGTGACAGTAGATAGGGTGGAGGCTGTGAGCAGGAGAGTGT 2880
QY 2881 TCTGTGGTGTGAGGGGTGTAGGGGAAAGCCAGACAGAGGAGTCTGGCTTGTGTCTCTGA 2940
Db 2881 TCTGTGGTGTGAGGGGTGTAGGGGAAAGCCAGACAGAGGAGTCTGGCTTGTGTCTCTGA 2940
QY 2941 CACATGTCTTACTTGTGTTATTAACAGGATGACCTGTGAAGACCAACATCTACGACTTC 3000
Db 2941 CACATGTCTTACTTGTGTTATTAACAGGATGACCTGTGAAGACCAACATCTACGACTTC 3000
QY 3001 TGAAGAAGCAGAGCCTTGAAGAGACAGGGGTGTCTCTGAGCCTTGGGCTTGAATGAGT 3060
Db 3001 TGAAGAAGCAGAGCCTTGAAGAGACAGGGGTGTCTCTGAGCCTTGGGCTTGAATGAGT 3060
QY 3061 CCACAAAGAGGAGCATGAGTGTGAGATTAAGGCCCAAGAGGCTTGAAGAGGCACTTG 3120
Db 3061 CCACAAAGAGGAGCATGAGTGTGAGATTAAGGCCCAAGAGGCTTGAAGAGGCACTTG 3120
QY 3121 GGAAGGGGTCAATCTGACAGAGCCCTTATTCATGGAATCTGAGAGCTTGGGGCCATCTG 3180
Db 3121 GGAAGGGGTCAATCTGACAGAGCCCTTATTCATGGAATCTGAGAGCTTGGGGCCATCTG 3180
QY 3181 TAAATCTGTGGGCTGGCAGGCAATTAAGGAGCAGCCTTCATCTGTGAGCAGCTGGGGA 3240
Db 3181 TAAATCTGTGGGCTGGCAGGCAATTAAGGAGCAGCCTTCATCTGTGAGCAGCTGGGGA 3240
QY 3241 GGGGGAAGGAGCAACCCCCCACTTAATCCCTTCTCCCTCAGCCCCAGAGTTAACACT 3300
Db 3241 GGGGGAAGGAGCAACCCCCCACTTAATCCCTTCTCCCTCAGCCCCAGAGTTAACACT 3300
QY 3301 CTGGCTTCCCTTCCCTTCCCACTTCCATCAGAGTGTGAAGGTTGCAAGGAGGAGTTAAA 3360
Db 3301 CTGGCTTCCCTTCCCTTCCCACTTCCATCAGAGTGTGAAGGTTGCAAGGAGGAGTTAAA 3360
QY 3361 CCTACATGTCCAAAATCATCTGAGTGAAGATATATGAGATCAGTATGTTAGAGCAAGAA 3420
Db 3361 CCTACATGTCCAAAATCATCTGAGTGAAGATATATGAGATCAGTATGTTAGAGCAAGAA 3420
QY 3421 GGAATCTGACAGGCTTAACTGGGTTAATGTGTAAGTCTGTGTGACATGTGTGTCTG 3480
Db 3421 GGAATCTGACAGGCTTAACTGGGTTAATGTGTAAGTCTGTGTGACATGTGTGTCTG 3480
QY 3481 ACTGAAAGCGGCAATGCTGTGACAGTGTTCAGTTCTGTGCTGAGCTTACAGACTGCA 3540
Db 3481 ACTGAAAGCGGCAATGCTGTGACAGTGTTCAGTTCTGTGCTGAGCTTACAGACTGCA 3540
QY 3541 GGTGTGTGTGTAATTTGCCAAGGCAAGTGGGTGAATCCCTTCAATGGTTTAAAGAGAT 3600
Db 3541 GGTGTGTGTGTAATTTGCCAAGGCAAGTGGGTGAATCCCTTCAATGGTTTAAAGAGAT 3600
QY 3601 TGAATGATGGCTGCACTTCAAGACCAATGAAAAATGAAATGACAATCTATATGTGTCT 3660
Db 3601 TGAATGATGGCTGCACTTCAAGACCAATGAAAAATGAAATGACAATCTATATGTGTCT 3660
QY 3661 CTGAGCTAAAGTGAAGGCTTTTGAAGACACTGTCTTGAAGATGTGGCAACAGAGAC 3720
Db 3661 CTGAGCTAAAGTGAAGGCTTTTGAAGACACTGTCTTGAAGATGTGGCAACAGAGAC 3720
QY 3721 TACAGACAGTATCTGTGACAGATTAAGAGAGAGAGGAGGGGTGTGAATTTCTTACTA 3780
Db 3721 TACAGACAGTATCTGTGACAGATTAAGAGAGAGAGGAGGGGTGTGAATTTCTTACTA 3780
QY 3781 TCAAAAGGAACTGAGTCTGTGACCTGCAAGTGAATGCTCTCCCTGACATCATGACTT 3840
Db 3781 TCAAAAGGAACTGAGTCTGTGACCTGCAAGTGAATGCTCTCCCTGACATCATGACTT 3840
QY 3841 TGTCTTGGGAGCCAGCACTGTGAACTTCAAGGTCTGAGAGATGAGAGGCTCCCTCA 3900
Db 3841 TGTCTTGGGAGCCAGCACTGTGAACTTCAAGGTCTGAGAGATGAGAGGCTCCCTCA 3900

QY 5567 GAGAAAGCAGCATTTCATGAGAGTCTGTTGGAGAGCCATAGAGCTACGGTGTAAAGA 5626
| | | | |
DB 1509 GAGAAAGCAGCATTTCATGAGAGTCTGTTGGAGAGCCATAGAGCTGTAAAGA 1568
| | | | |
QY 5627 GGCAGGGAAGTGTGTGTGAGAAAGTCAGAGCTTCATGAAAGCTTACCCACACAG 5686
| | | | |
DB 1569 GGCAGGGAAGTGTGTGTGAGAAAGTCAGAGCTTCATGAAAGCTTACCCACACAG 1628
| | | | |
QY 5687 AAATGACAGACAGATCCCTCTATCTCCCATTAAGATTGAGT 5731
| | | | |
DB 1629 AAATGACAGACAGATCCCTCTATCTCCCATTAAGATTGAGT 1673
| | | | |
RESULT 5
ABV74357
ID ABV74357 standard; DNA, 1679 BP.
AC ABV74357,
XX
XX 05-FEB-2003 (first entry)
DT
XX Mouse cardiac alpha-myosin heavy chain promoter fragment.
DE
XX Therapeutic delivery system; electrical pulse generator; brain;
KM electrically responsive promoter; ischaemic injury; cardiac; kidney;
KM arterial occlusive disease; coronary arterial disease; stroke; mouse;
KM alpha myosin; promoter; de.
XX
OS Mus BP.
XX WO200249669-A2.
PN
XX 27-JUN-2002.
PD
XX 20-DEC-2001; 2001WO-US049845.
PF
XX 21-DEC-2000; 2000US-0257460P.
PR 20-AUG-2001; 2001US-0313926P.
XX
XX (MEDT) MEDTRONIC INC.
PA
XX Schu CA, Padua R, Bonner M, Donovan MG, Soykan O;
PI
XX WPI: 2002-537680/57.
DR
XX Therapeutic delivery system used for treating e.g. peripheral arterial
PT occlusive disease or stroke comprises electrical pulse generator operably
PT coupled with genetically engineered cells in mammalian tissue.
PT
XX Disclosure; Page: 66pp + Sequence Listing; English.
XX
XX The invention relates to a therapeutic delivery system comprising an
CC electrical pulse generator operably coupled with genetically engineered
CC cells in a mammalian tissue. The genetically engineered cells also
CC comprise a target gene operably coupled to an electrically responsive
CC promoter. The delivery system is useful for the production of
CC therapeutically useful gene or protein products, in repairing tissue
CC injury (e.g. ischaemic injury, damaged cardiac tissue, kidney tissue,
CC brain tissue or endothelial tissue), in stimulating cells for controlled
CC expression of therapeutically useful gene and protein sequences and for
CC treating peripheral arterial occlusive disease, coronary arterial disease
CC or stroke. The present sequence is that of the mouse cardiac alpha-myosin
CC heavy chain promoter fragment, exemplary of cardiac specific promoter
CC regions. Note: The sequence data for this patent is not represented in
CC the printed specification but is based on sequence information supplied
CC to Derwent by the European Patent Office
CC
XX Sequence 1679 BP; 342 A; 486 C; 418 G; 433 T; 0 U; 0 Other;
SQ

Query Match 22.3%; Score 1279.8; DB 6; Length 1679;
Best Local Similarity 82.2%; Pred. No. 0; Mismatches 57; Indels 292; Gaps 1;
Matches 1616; Conservative 0;

QY 3767 GAATTCCTTACTATCAAAAGGAAACCTGAGCTGACCTGCAAAAGTGTCTCCCT 3826
| | | | |
DB 1 GAATTCCTTACTATCAAAAGGAAACCTGAGCTGACCTGCAAAAGTGTCTCCCT 60
| | | | |
QY 3827 AGACATCATGACTTTGTCTCTGGGAGCCAGACACTGTGAACTTCAGGCTGAGAGAGTA 3886
| | | | |
DB 61 AGACATCATGACTTTGTCTCTGGGAGCCAGACACTGTGAACTTCAGGCTGAGAGAGTA 120
| | | | |
QY 3887 GGAAGCTCCCTCAGCCTGAAAGTATGAGATAGCAGGAGTTGAAAGGGGAAAGGAGAG 3946
| | | | |
DB 121 GGAAGCTCCCTCAGCCTGAAAGTATGAGATAGCAGGAGTTGAAAGGGGAAAGGAGAG 180
| | | | |
QY 3947 CCTGGAGATGAGACTTGTGTGTTGAGAGGAGGAGCAGATATTAAAGCTTGAAGAGAG 4006
| | | | |
DB 181 CCTGGAGATGAGACTTGTGTGTTGAGAGGAGGAGCAGATATTAAAGCTTGAAGAGAG 240
| | | | |
QY 4007 TGACCTTTACCCAGCTTGTTCACACTCACCTTCAGATTTAAATAATTAAGGTAAAGGCTT 4066
| | | | |
DB 241 TGACCTTTACCCAGCTTGTTCACACTCACCTTCAGATTTAAATAATTAAGGTAAAGGCTT 300
| | | | |
QY 4067 GGGTAGGGAGGTGTGTGAGAGCGCTCTGTCTCTCTGATGCTGAGGCTTGAAGGCTTTG 4126
| | | | |
DB 301 GGGTAGGGAGGTGTGTGAGAGCGCTCTGTCTCTCTGATGCTGAGGCTTGAAGGCTTTG 360
| | | | |
QY 4127 GGGAGAGGAAATGTGCCCAGAGACTTAAATAAAAGCCATGAGCAGAGGGGCGAGGCA 4186
| | | | |
DB 361 GGGAGAGGAAATGTGCCCAGAGACTTAAATAAAAGCCATGAGCAGAGGGGCGAGGCA 420
| | | | |
QY 4187 CAGACCTTTCTATGAGGAAACCTTGGGGCCCTGATGATGATTGAACAAGACTCGCCAT 4246
| | | | |
DB 421 CAGACCTTTCTATGAGGAAACCTTGGGGCCCTGATGATGATTGAACAAGACTCGCCAT 480
| | | | |
QY 4247 CGATACCTTCTCTCTTAACGAGCAGAGAGGAACCTGAGTTTACCACTCCCTATCAGTG 4306
| | | | |
DB 481 CGATACCTTCTCTCTTAACGAGCAGAGAGGAACCTGAGTTTACCACTCCCTATCAGTG 525
| | | | |
QY 4307 ATGAGAGAAAGTGAAGTGAATTTCACCTCCCTATCAGTATGAGAAAGTGAAGT 4366
| | | | |
DB 526 ATGAGAGAAAGTGAAGTGAATTTCACCTCCCTATCAGTATGAGAAAGTGAAGT 525
| | | | |
QY 4367 CGAGTTTACCACTCCCTATCAGTATGAGAAAGTGAAGTGAAGTGAAGTGAAGT 4426
| | | | |
DB 526 CGAGTTTACCACTCCCTATCAGTATGAGAAAGTGAAGTGAAGTGAAGTGAAGT 525
| | | | |
QY 4427 TCAGTATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4486
| | | | |
DB 526 TCAGTATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 525
| | | | |
QY 4487 GAAAGTCGAGTTTACCACTCCCTATCAGTATGAGAAAGTGAAGTGAAGTGAAGT 4546
| | | | |
DB 526 GAAAGTCGAGTTTACCACTCCCTATCAGTATGAGAAAGTGAAGTGAAGTGAAGT 525
| | | | |
QY 4547 TCCCTTACAGTATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 4606
| | | | |
DB 526 TCCCTTACAGTATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 548
| | | | |
QY 4607 TTAGAGCAGCAGCATTTGGATGAGATATTAAGAGGCTGAGAGCACTGAGAGTGAAG 4666
| | | | |
DB 549 TTAGAGCAGCAGCATTTGGATGAGATATTAAGAGGCTGAGAGCACTGAGAGTGAAG 608
| | | | |
QY 4667 ATTTCTTCAACCCAGGTAAGAGGAGTTCGGGTGGGGCTCTTCAACCCACACAGACT 4726
| | | | |
DB 609 ATTTCTTCAACCCAGGTAAGAGGAGTTCGGGTGGGGCTCTTCAACCCACACAGACT 668
| | | | |
QY 4727 CTCGCCACCTTGAAGAAATGCTTCTTCTGAGAGTGGGGTTGAGGCTGAGAGTCT 4786
| | | | |
DB 669 CTCGCCACCTTGAAGAAATGCTTCTTCTGAGAGTGGGGTTGAGGCTGAGAGTCT 728
| | | | |
QY 4787 GACAGGGTGGCTTTCACACAGCCTGAGAGTTCCTCAGTGGCAGAGGTTTCCACAAGAA 4846
| | | | |
DB 729 GACAGGGTGGCTTTCACACAGCCTGAGAGTTCCTCAGTGGCAGAGGTTTCCACAAGAA 788
| | | | |

QY	4847	CACGGAATGCCCTCCCTTCAAGCGATCTTCCATCTCTCCGAGGAGATGCTCC	4906
Db	789	CACGAGATGCCCTCCCTTCAAGCGATCTTCCATCTCTCCGAGGAGATGCTCC	848
QY	4307	CGCTTGATTAATCTTGAGCTCTTGCTGATTCAGCAAGATTTTGCCTGATGCACTCA	4966
Db	849	CGCTTGATTAATCTTGAGCTCTTGCTGATTCAGCAAGATTTTGCCTGATGCACTCA	908
QY	4367	TCCTTCTCTACTGCTCCGAGGCTTGCGTTCGCTTCTTGCGATGCTCTTTCACCCA	5026
Db	909	TCCTTCTCTACTGCTCCGAGGCTTGCGTTCGCTTCTTGCGATGCTCTTTCACCCA	968
QY	5027	TTTCTCACTTCACCTTTTCTCCCTTCATTTGATATCATCTCTCTCTCTCTCT	5086
Db	969	TTTCTCACTTCACCTTTTCTCCCTTCATTTGATATCATCTCTCTCTCTCTCTCT	1028
QY	5087	TCCTTCT	5146
Db	1029	TCCTTCT	1088
QY	5147	TCCTTCT	5206
Db	1089	TCCTTCT	1148
QY	5207	TAAACAATCTTCCAGTGAAGCAACAGCTTCAGTGTCTGGGTGCTCTTAACTTCTCA	5266
Db	1149	TAAACAATCTTCCAGTGAAGCAACAGCTTCAGTGTCTGGGTGCTCTTAACTTCTCA	1208
QY	5267	CCCCCTGGCTTGTCTGTTCATCTGGTCAAGAACTCTAATATGTCTCCACAGCTCTG	5326
Db	1209	CCCCCTGGCTTGTCTGTTCATCTGGTCAAGAACTCTAATATGTCTCCACAGCTCTG	1268
QY	5327	CTACTCTCTTCTCTGCGCTGTCTCTCTGTCCAGCTGCGCACTGTGTGCTCTGTTC	5386
Db	1269	CTACTCTCTTCTCTGCGCTGTCTCTCTGTCCAGCTGCGCACTGTGTGCTCTGTTC	1328
QY	5387	CAGCTGTGTCTCAATCTTTCAGGATTTCTCTGAAAAGTTTAACTCAAGTGAAGATTTTCCC	5446
Db	1329	CAGCTGTGTCTCAATCTTTCAGGATTTCTCTGAAAAGTTTAACTCAAGTGAAGATTTTCCC	1388
QY	5447	CTGTAGACAGAGATTCAGATTTCTCCGGAAGTCAAGCTTCCAGCCCTCTCTTCTCTGC	5506
Db	1389	CTGTAGACAGAGATTCAGATTTCTCCGGAAGTCAAGCTTCCAGCCCTCTCTTCTCTGC	1448
QY	5507	CCAGCTGCCCGGCACTCTTTCAGCAAACTCAGGCAACCTTACCCCAATGACTCTTGACA	5566
Db	1449	CCAGCTGCCCGGCACTCTTTCAGCAAACTCAGGCAACCTTACCCCAATGACTCTTGACA	1508
QY	5567	GAGAAAGCAGGACCTTTCATGAGTCCGTGTGGGAGAGCCATTAAGCTTACGGTGTAAAGA	5626
Db	1509	GAGAAAGCAGGACCTTTCATGAGTCCGTGTGGGAGAGCCATTAAGCTTACGGTGTAAAGA	1568
QY	5627	GGCAGGGAAGTGTGTGTAGAAAAGTCAGACTTCACATTAAGAAAGCTTACGCCACACAG	5686
Db	1569	GGCAGGGAAGTGTGTGTAGAAAAGTCAGACTTCACATTAAGAAAGCTTACGCCACACAG	1628
QY	5687	AAATGACAGACAGATCCCTCTATCTCCCCCATTAAGATTTGAAT 5731	
Db	1629	AAATGACAGACAGATCCCTCTATCTCCCCCATTAAGATTTGAAT 1673	

RESULT 6
ADL14241
ID ADL14241 standard; DNA; 1679 BP.
XX
AC ADL14241;
XX
DT 06-MAY-2004 (first entry)
XX
DE Mouse cardiac alpha-myosin heavy chain promoter region.
XX
KW Mouse; cardiac alpha-myosin heavy chain promoter; ds;
KW electrical pulse generator; electrically responsive promoter;
KW electrical

KW	electrical response enhancer element; pacemaker.
XX	
OS	Mus sp.
XX	
PN	US2003204206-A1.
XX	
PD	30-OCT-2003.
XX	
PF	20-DEC-2001; 2001US-00027655.
XX	
PR	21-DEC-2000; 2000US-0257460P.
PR	20-AUG-2001; 2001US-0313926P.
XX	
PA	(MEDT) MEDTRONIC INC.
XX	
PI	Padua RA, Schu CA, Bonner MD, Donovan MG, Soykan O;
XX	
DR	WPI; 2004-032680/03.
XX	
PT	Therapeutic delivery system useful for regulating delivery of therapeutic
PT	proteins and nucleic acids, comprises electrical pulse generator coupled
PT	with genetically engineered cells in mammalian tissue.
XX	
PS	Disclosure; SEQ ID NO 5; 39pp; English.

CC The invention relates to a therapeutic delivery system comprising an
 CC electrical pulse generator coupled with genetically engineered cells in
 CC mammalian tissue. The genetically engineered cells further include a
 CC target gene coupled to an electrically responsive promoter. The invention
 CC also relates to an expression vector comprising an electrical response
 CC enhancer element, a tissue specific promoter heterologous to the element
 CC and a coding sequence, an apparatus for testing cells comprising an upper
 CC plate electrode, a lower plate electrode and a porous membrane positioned
 CC between electrodes during operation, and a method of treating a patient
 CC comprising providing the patient with an electrical pulse generator
 CC coupled with genetically engineered cells in a patient tissue. The
 CC electrical pulse generator is a pacemaker. The method is used for
 CC regulating the delivery of therapeutic proteins and nucleic acids. The
 CC invention provides controlled and local delivery of therapeutically
 CC important gene or protein products. This sequence represents the mouse
 CC cardiac alpha-myosin heavy chain promoter region, used in the method of
 CC the invention.
 XX
 SQ Sequence 1679 BP; 342 A; 486 C; 418 G; 433 T; 0 U; 0 Other;

Query Match	22.3%;	Score 1279.8;	DB 12;	Length 1679;
Best Local Similarity	82.2%;	Pred. No. 0;		
Matches 1616; Conservative	0;	Mismatches 57;	Indels 292;	Gaps 1;

QY	376	AAATTTCCTTA	CTATCAAAAGG	AAACGTAGTCG	GCACCTCG	CAAAAGTGAATGC	CTCCCT	3826	
Db	1	GAATTTCTTTA	CTATCAAAAGG	AAACGTAGTCG	GCACCTCG	CAAAAGTGAATGC	CTCCCT	60	
QY	3827	AGACATCA	TGACTTTGTCTCT	GTGGGAGCC	AGACACTGTG	GAACCTTCAG	GTCTGAGAGTA	3886	
Db	61	AGACATCA	TGACTTTGTCTCT	GTGGGAGCC	AGACACTGTG	GAACCTTCAG	GTCTGAGAGTA	120	
QY	3887	GGAGGCTCC	CTCAGCCTG	AAAGCTATG	CAATGATCC	AGGGTTG	GAAGGGGGAA	GGAGAG	3946
Db	121	GGAGGCTCC	CTCAGCCTG	AAAGCTATG	CAATGATCC	AGGGTTG	GAAGGGGGAA	GGAGAG	180
QY	3947	CCTGGATGG	AGCTTTGTGTGTT	GAAGGCA	GGGACAGAT	TTAAGCTTG	GAAGAGAGG	4006	
Db	181	CCTGGATGG	AGCTTTGTGTGTT	GAAGGCA	GGGACAGAT	TTAAGCTTG	GAAGAGAGG	240	
QY	4007	TGACCCCTTA	CCAGATGTTCA	CTACCCCTT	CAGATTAAAA	ATACTAG	GGTAAGG	4066	
Db	241	TGACCCCTTA	CCAGATGTTCA	CTACCCCTT	CAGATTAAAA	ATACTAG	GGTAAGG	300	
QY	4067	GGGTAAGG	GAAGTGTGTG	AGAGCCTCT	GTCTCTCTCT	CGCATG	CCCTGAAGG	4126	
Db	301	GGGTAAGG	GAAGTGTGTG	AGAGCCTCT	GTCTCTCTCT	CGCATG	CCCTGAAGG	360	

PT Nucleic acid encoding tetracycline-inducible transcription regulatory
 PT fusion protein - comprising modified tetracycline repressor able to bind
 PT mutant tet operator, fused to transcription regulator, useful for
 PT modulating eukaryotic gene expression.
 PS Disclosure; Page 79-80; 117pp; English.
 XX
 CC AAT45721 and AAT45722 are bi-directional, tetracycline (Tc)-regulated
 CC promoters used to co-ordinate regulation of expression of 2 genes of
 CC interest. The promoters contain tet operator sequences and a minimal
 CC herpes simplex virus (HSV) thymidine kinase (tk) promoter. The main
 CC invention of the specification concerns modified Tet repressor (Tetr)
 CC proteins that bind to modified class B tet operator sequences tetO-4C and
 CC tetO-6C (see AAT45711 and AAT48478). Modified Tetr proteins can be fused
 CC to any transcription regulatory polypeptide (e.g. HSV virion protein 16,
 CC VP16) and used to control transcription of a tetO-4C or tetO-6C linked
 CC gene. Nucleic acid encoding such a fusion protein may be introduced into
 CC a cell and transcription of the protein can be controlled by altering the
 CC concn. of tetracycline (or an analogue) in the cell, as appropriate. This
 CC ability to modulate gene expression in a predictable way is very useful
 CC in gene therapy and for recombinant protein prodn. in cultured cells or
 CC transgenic animals. The Tc-inducible system is also useful for the prodn.
 CC of transgenic animal models for the study of disease and also for the
 CC study of gene function e.g. during differentiation. The Tc-inducible
 CC system allows rapid activation of gene transcription without cellular
 CC toxicity, high concns. of inducer are not required
 XX
 SQ Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;
 XX
 Query Match 5.4%; Score 311.8; DB 2; Length 520;
 Best Local Similarity 99.4%; Pred. No. 8.7e-78;
 Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4277 GGAATCGAATTACCACTCCCTATCAGTATGAGAAAAGTGAAGTGAATTACCA 4336
 DB 63 GGATCTCGAGTTTACCACTCCCTATCAGTATGAGAAAAGTGAAGTGAATTACCA 122
 QY 4337 TCCCTATCAGTATGAGAAAAGTGAAGTGAATTACCACTCCCTATCAGTATGAG 4396
 DB 123 TCCCTATCAGTATGAGAAAAGTGAAGTGAATTACCACTCCCTATCAGTATGAG 182
 QY 4397 AAAAGTGAAGTGAAGTGAATTACCACTCCCTATCAGTATGAGAAAAGTGAAGT 4456
 DB 183 AAAAGTGAAGTGAAGTGAATTACCACTCCCTATCAGTATGAGAAAAGTGAAGT 242
 QY 4457 TACCACTCCCTATCAGTATGAGAAAAGTGAAGTGAATTACCACTCCCTATCAGTG 4516
 DB 243 TACCACTCCCTATCAGTATGAGAAAAGTGAAGTGAATTACCACTCCCTATCAGTG 302
 QY 4517 ATAGAGAAAAGTGAAGTGAATTACCACTCCCTATCAGTATGAGAAAAGTGAAGT 4576
 DB 303 ATAGAGAAAAGTGAAGTGAATTACCACTCCCTATCAGTATGAGAAAAGTGAAGT 362
 QY 4577 CGAGCTCGGTACGAG 4591
 DB 363 CGAGCTCGGTACCG 377
 XX
 RESULT 10
 AAV60079
 ID AAV60079 standard; DNA; 520 BP.
 XX
 AC AAV60079;
 XX
 DT 25-MAR-2003 (revised)
 DT 04-DEC-1998 (first entry)
 XX
 DE Bidirectional promoter region for coordinate regulation of 2 genes.
 XX
 XX Tet repressor; tetracycline; regulation; expression;
 KM Tet operator-linked gene; bidirectional promoter region;
 KM coordinate regulation; tetracycline-regulated transcriptional activator;
 KM db.

XX
 OS Synthetic.
 XX
 PN US5814618-A.
 XX
 XX 29-SEP-1998.
 PD
 PF 07-JUN-1995; 95US-00485978.
 XX
 PR 14-JUN-1993; 93US-00076327.
 PR 14-JUN-1993; 93US-00076726.
 PR 14-JUN-1994; 94US-00260452.
 PR 01-JUL-1994; 94US-00270637.
 PR 15-JUL-1994; 94US-00275876.
 PR 03-FEB-1995; 95US-00383754.
 XX
 PA (KNOL) KNOL AG.
 PA (BADI) BASF AG.
 XX
 PI Gossen M, Bujard H;
 XX
 DR WPI; 1998-541795/46.
 XX
 PT Tetracycline based regulation of gene expression - uses a tetracycline
 PT operator sequence joined to a gene of interest, the gene of interest
 PT being induced in the presence, but not absence of the antibiotic.
 XX
 PS Disclosure; Fig 7B; 63pp; English.
 XX
 SQ The present sequence represents a bidirectional promoter region for
 CC coordinate regulation of 2 genes by a tetracycline-regulated
 CC transcriptional activator. The specification describes a method for
 CC regulating expression of a Tet (tetracycline) operator-linked gene in a
 CC cell of a subject. The method comprises introducing into the cell a
 CC nucleic acid encoding a fusion protein which inhibits transcription in
 CC eukaryotic cells, the fusion protein comprising a polypeptide which binds
 CC to a Tet operator sequence, operatively linked to heterologous second
 CC polypeptide which inhibits transcription in eukaryotic cells and
 CC modulating the concentration of a tetracycline (analogue) in the subject.
 CC The method is used for the regulation of gene expression system, using
 CC tetracycline (analogues). The system enables a gene coupled to the system
 CC to be induced in the presence of Tet and then stopped when Tet is
 CC removed. (Updated on 25-MAR-2003 to correct PR field.)
 XX
 SQ Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;
 XX
 Query Match 5.4%; Score 311.8; DB 2; Length 520;
 Best Local Similarity 99.4%; Pred. No. 8.7e-78;
 Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 4277 GGAATCGAATTACCACTCCCTATCAGTATGAGAAAAGTGAAGTGAATTACCA 4336
 DB 63 GGATCTCGAGTTTACCACTCCCTATCAGTATGAGAAAAGTGAAGTGAATTACCA 122
 QY 4337 TCCCTATCAGTATGAGAAAAGTGAAGTGAATTACCACTCCCTATCAGTATGAG 4396
 DB 123 TCCCTATCAGTATGAGAAAAGTGAAGTGAATTACCACTCCCTATCAGTATGAG 182
 QY 4397 AAAAGTGAAGTGAAGTGAATTACCACTCCCTATCAGTATGAGAAAAGTGAAGT 4456
 DB 183 AAAAGTGAAGTGAAGTGAATTACCACTCCCTATCAGTATGAGAAAAGTGAAGT 242
 QY 4457 TACCACTCCCTATCAGTATGAGAAAAGTGAAGTGAATTACCACTCCCTATCAGTG 4516
 DB 243 TACCACTCCCTATCAGTATGAGAAAAGTGAAGTGAATTACCACTCCCTATCAGTG 302
 QY 4517 ATAGAGAAAAGTGAAGTGAATTACCACTCCCTATCAGTATGAGAAAAGTGAAGT 4576
 DB 303 ATAGAGAAAAGTGAAGTGAATTACCACTCCCTATCAGTATGAGAAAAGTGAAGT 362
 QY 4577 CGAGCTCGGTACGAG 4591
 DB 363 CGAGCTCGGTACCG 377

RESULT 11

AA60045 ID AAX60045 standard; DNA; 520 BP.

AC AAX60045;

DT 20-MAR-2003 (revised)
DT 04-AUG-1999 (first entry)

DE Bidirectional promoter construct for regulation of 2 genes.

XX Transgenic mice; transgene; tet operator-linked gene; tetracycline;
 KM mouse-actve transcriptional regulatory element; mutant Tet repressor;
 KM gene therapy; genetic disease; acquired disease; cancer; viral disease;
 KM vaccination; rheumatoid arthritis; hypopituitarism; wound healing;
 KM tissue regeneration; cancer; benign prostatic hypertrophy; hemophilia;
 KM erythrocytopenia; atherosclerosis; liver disease; Alzheimer's disease;
 KM Parkinson's disease; human disease model; ds.

XX Unidentified.

XX US5912411-A.

XX 15-JUN-1999.

XX 07-JUN-1995; 95US-00487472.

XX 14-JUN-1993; 93US-00076327.

XX 14-JUN-1993; 93US-00076726.

XX 14-JUN-1994; 94US-00260452.

XX 01-JUL-1994; 94US-00270637.

XX 15-JUL-1994; 94US-00275876.

XX 03-FEB-1995; 95US-00383754.

XX (UYHE-) UNIV HEIDELBERG.

XX Bujard H, Gossen M;

XX WPI; 1999-357232/30.

XX Transgenic mice with inducible transgene activity useful for in vitro and

XX in vivo protein production.

XX Disclosure; Fig 7B; 63pp; English.

XX The specification describes transgenic mice which have a transgene and a
 CC tet operator-linked gene integrated in the genome. The transgene
 CC comprises a mouse-actve transcriptional regulatory element linked to a
 CC polynucleotide sequence that encodes a fusion protein which activates
 CC transcription of the tet operator-linked gene. The fusion protein
 CC comprises a mutated Tet repressor that binds a tet operator sequence in
 CC the presence of tetracycline (Tc) or a Tc analogue, linked to a
 CC polypeptide that activates transcription in eukaryotic cells. The
 CC transgenic system may be used for gene therapy to treat genes involved in
 CC genetic or acquired diseases. Gene therapy may be used to treat cancer,
 CC viral diseases, for vaccination, and to provide (Tc induced) regulated
 CC doses of a product (e.g. for the treatment or regulation of rheumatoid
 CC arthritis, hypopituitarism, wound healing and tissue regeneration,
 CC cancer, benign prostatic hypertrophy, hemophilia, erythrocytopenia,
 CC atherosclerosis and liver disease, Alzheimer's disease, and Parkinson's
 CC disease). The system may also be used to produce proteins in vivo (e.g.
 CC using mammalian, yeast or fungal cells) or in vitro (e.g. transgenic farm
 CC animals), to produce animal models of human disease, or to produce a
 CC stable cell line for gene cloning. The present sequence represents a
 CC bidirectional promoter construct used to control the regulation of two
 CC genes by a Tc-regulated transcriptional activator. (Updated on 20-MAR-
 CC 2003 to correct PF field.)

XX Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;

XX Query Match 5.4%; Score 311.8; DB 2; Length 520;

Best Local Similarity 99.4%; Pred. No. 8.7e-78;
 Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 4277 GGAACCTCGAGTTTACACCTCCCTATCATGATAGAGAAAGTGAAGTGAAGTTTACAC 4336
 Db 63 GGAATCTGAGTTTACACCTCCCTATCATGATAGAGAAAGTGAAGTGAAGTTTACAC 122
 QY 4337 TCCCTATCATGATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4396
 Db 123 TCCCTATCATGATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 182
 QY 4397 AAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4456
 Db 183 AAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 242
 QY 4457 TACCACTCCCTATCATGATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4516
 Db 243 TACCACTCCCTATCATGATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 302
 QY 4517 ATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 4576
 Db 303 ATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 362
 QY 4577 CGAGCTCGGTACCG 4591
 Db 363 CGAGCTCGGTACCG 377

RESULT 12

AAZ56126 ID AAZ56126 standard; DNA; 520 BP.

XX AAZ56126;

XX 27-MAR-2000 (first entry)

XX Bidirectional promoter region derived from CMV promoter.

XX Tetracycline; tetracycline repressor; gene expression regulation; cancer;

XX gene therapy; arthritis; wound healing; tissue regeneration; promoter;

XX ds.

XX Cytomegalovirus.

XX US6004941-A.

XX 21-DEC-1999.

XX 07-JUN-1995; 95US-00485740.

XX 14-JUN-1993; 93US-00076327.

XX 14-JUN-1993; 93US-00076726.

XX 14-JUN-1994; 94US-00260452.

XX 01-JUL-1994; 94US-00270637.

XX 15-JUL-1994; 94US-00275876.

XX 03-FEB-1995; 95US-00383754.

XX (BADI) BASF AG.

XX (BADI) BASF BIORESEARCH CORP.

XX (KNOL) KNOLL AG.

XX Gossen M, Bujard H;

XX WPI; 2000-085798/07.

XX Regulation of gene expression in cells, useful for gene therapy of

XX diseases, production of proteins in vitro and in vivo and production of

XX stable cell lines for cloning.

XX Disclosure; Fig 7B; 64pp; English.

XX This sequence represents a bidirectional promoter, for use in the method

XX of the invention. The invention relates to a method for regulation gene

PR 15-JUN-1994; 94US-00275876.
PR 03-FEB-1995; 95US-00383754.
PR 07-JUN-1995; 95US-00485978.
PR 28-SEP-1998; 98US-00162184.
XX
PA (BADI) BASF AG.
PA (KNOL) KNOLL AG.
XX
PL Bujard H, Gossen M,
XX
DR MPI; 2001-556625/62.
XX
PT Fusion protein for inhibiting transcription in eukaryotic cells useful in
PT gene therapy applications comprises a first polypeptide, which binds to
PT tel operator sequences, operatively linked to a heterologous second
PT polypeptide.
XX

XX

CC The invention relates to a fusion protein that comprises a first
CC polypeptide which binds to tet operator sequences, operatively linked to

CC a heterologous second polypeptide, which inhibits transcription in
CC eukaryotic cells. The fusion proteins are tetracycline-responsive and are
CC useful for regulation of transcription in eukaryotic cells and animals.
CC The tetracycline (Tc)-controlled regulatory system is useful in various

CC applications in gene therapy, such as in the treatment of various disease
CC conditions e.g. rheumatoid arthritis, hypopituitarism, wound healing and
CC tissue regeneration, anticancer treatments, benign prostatic hypertrophy,
CC hemophilia, diabetes and atherosclerosis. They are also useful for bone

CC marrow support therapy, treatment of central nervous system disorders
CC e.g. Alzheimer's disease, Parkinson's disease (see AH47628 for a
CC detailed description of the uses). The present sequence represents the
CC nucleotide sequence of a bidirectional promoter region for coordinate

CC regulation of two genes or interest by tetracycline-regulated transcriptional activator

XX Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;

Query Match	5.4%	Score 311.8;	DB 4;	Length 520;
Best Local Similarity	99.4%;	Pred. No. 8,7e-78;		
Matches 313; Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0;

OY 4277 GGAAGCTCGAGTTTACCACTCCCTATCAGTGATAGAAAAAGTGAAGTCGAGTTTAAACA 4336
|||||
Db 63 GGAAGCTCGAGTTTACCACTCCCTATCAGTGATAGAAAAAGTGAAGTCGAGTTTAAACA 122

Oy	4337	TCCTATCAGTATAGAGAAAGTGAAGTTCACCTCCCTATCATGTATGAG	4336
Db	123	TCCTATCAGTATAGAGAAAGTGAAGTTCACCTCCCTATCATGTATGAG	182

Qy	4397	AAAGTGAAGTCAGATTACATCTCCCTTACAGTGTATGAGAAAGTGAAGTCGAGTT	4456
Db	183	AAAGTGAAGTCAGTTACCATCTCCCTTACAGTGTATGAGAAAGTGAAGTCGAGTT	242

Accession	Sequence	Position
Qy	TACCACTCCCTATAGATAGAAAAAGTGAAGTTCAGTTTACACCTCCCTATCAGTG	4516
Db	TACCACTCCCTATCAGTATGAGAAAAAGTGAAGTTCAGTTTACCACTCCCTATCAGTG	302

Accession	Sequence	Length
Dy	ATTAGAAAATGTGAAGTCGAGTTTACCACTCCCTTACAGTAGAGAAAAGTGAAGT	45/16
Db	ATTAGAAAATGTGAAGTCGAGTTTACCACTCCCTTACAGTAGAGAAAAGTGAAGT	36/2

[illegible]

RESULT 14
AAH25572
ID AAH25572 standard; DNA; 520 BP.

XX
AC AAH25572;
XX

DT 05-SEP-2001 (first entry)
XX Nucleotide sequence of a bidirectional promoter.
DE
XX Transgenic plant; transgene; tet operator-linked gene; Tet repressor;
KM tetracycline-regulated transcriptional regulatory system; tet operator;
XX tetracycline; ds.
XX
OS Unidentified.
PN US6242667-B1.
PD 05-JUN-2001.
XX 28-SEP-1998; 98US-00161902.
PF 14-JUN-1993; 93US-00076327.
PR 14-JUN-1993; 93US-00076726.
PR 14-JUN-1994; 94US-00260452.
PR 01-JUL-1994; 94US-00270637.
PR 15-JUL-1994; 94US-00275876.
PR 03-FEB-1995; 95US-00383754.
PR 07-JUN-1995; 95US-00487472.
XX
PA (BAD1) BASF AG.
PA (KNOL) KNOLL AG.
PI Bujard H, Gossen M;
XX WPI; 2001-396837/42.
DR
XX Transgenic plant for regulating the expression of genes, comprises a
PT transgene integrated into the genome and a tet operator-linked gene in
PT the genome.
XX
PS Disclosure; Fig 7B; 65pp; English.
XX
XX The specification describes a transgenic plant with a tetracycline-
CC regulated transcriptional regulatory system. The transgenic plants have a
CC transgene and a tet operator-linked gene integrated in the genome, which
CC confers a detectable and functional phenotype on the plant when
CC expressed. The transgene comprises a transcriptional regulatory element
CC functional in cells of the plant operatively linked to a polynucleotide
CC sequence encoding a fusion protein that activates transcription of the
CC tet operator linked gene. The fusion protein comprises a first
CC polypeptide that is a mutated Tet repressor that binds to a tet operator
CC sequence in the presence of tetracycline or its analogue, operatively
CC linked to a second polypeptide that activates transcription in eukaryotic
CC cells. In the presence of tetracycline, the fusion protein binds to the
CC tet operator-linked gene and activates transcription of the tet operator-
CC linked gene such that it is expressed at a level sufficient to confer the
CC detectable and functional phenotype on the plant. The level of expression
CC of the tet-operator gene can be downmodulated by depleting tetracycline
CC from the plant. The transgene can be used to regulate the expression of
CC genes in the plant. The transgenic plant can be used to analyse the
CC functions of cellular proteins. The present sequence represents a
CC bidirectional promoter for coordinate regulation of two genes of interest
CC by a tetracycline-regulated transcriptional activator. It is used to
CC produce transgenic plants of the invention
XX
SQ Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;
Query Match 5.4%; Score 311.8; DB 5; Length 520;
Best Local Similarity 99.4%; Pred. No. 8-7e-78;
Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 4277 GGAAGTCAAGTTTACCTCCCTATCAGTGTATAGAAAAGTGAAGTCAAGTTTACAC 4336
DB 63 GGATCTCGAGTTTACCTCCCTATCAGTGTATAGAAAAGTGAAGTCAAGTTTACAC 122
QY 4337 TCCCTATCAGTGTATAGAAAAGTGAAGTCAAGTTTACCTCCCTATCAGTGTATAG 4396
DB 123 TCCCTATCAGTGTATAGAAAAGTGAAGTCAAGTTTACCTCCCTATCAGTGTATAG 182

QY 4397 AAAAGTGAAGTCAAGTTTACCTCCCTATCAGTGTATAGAAAAGTGAAGTCAAGTT 4456
DB 183 AAAAGTGAAGTCAAGTTTACCTCCCTATCAGTGTATAGAAAAGTGAAGTCAAGTT 242
QY 4457 TACCACTCCCTATCAGTGTATAGAAAAGTGAAGTCAAGTTTACCTCCCTATCAGTG 4516
DB 243 TACCACTCCCTATCAGTGTATAGAAAAGTGAAGTCAAGTTTACCTCCCTATCAGTG 302
QY 4517 ATAGAGAAAAGTGAAGTCAAGTTTACCTCCCTATCAGTGTATAGAAAAGTGAAGT 4576
DB 303 ATAGAGAAAAGTGAAGTCAAGTTTACCTCCCTATCAGTGTATAGAAAAGTGAAGT 362
QY 4577 CGAGCTCGGTACCG 4591
DB 363 CGAGCTCGGTACCG 377
RESULT 15
ACA94734
ID ACA94734 standard; DNA; 520 BP.
XX
AC ACA94734;
XX
DT 11-AUG-2003 (first entry)
XX
XX Tet coordinated bi-directional promoter region #2.
DE
XX Gene therapy; tet promoter; transgenic; rheumatoid arthritis; ds;
KM hypodermatism; wound healing; anti-cancer treatment; promoter;
XX transgenic farm animal; stable cell line production; tetracycline.
OS
XX Unidentified.
XX
PN US2003022315-A1.
PD 30-JAN-2003.
PF 03-AUG-2001; 2001US-00921650.
XX
PR 14-JUN-1993; 93US-00076327.
PR 01-JUL-1994; 94US-00270637.
PR 15-JUL-1994; 94US-00275876.
PR 03-FEB-1995; 95US-00383754.
PR 07-JUN-1995; 95US-00485978.
PR 28-SEP-1998; 98US-00162184.
PR 24-JAN-2000; 2000US-00489777.
XX
PA (BAD1) BASF AG.
PA (KNOL) KNOLL AG.
PI Bujard H, Gossen M;
XX WPI; 2003-438975/41.
DR
XX Regulating expression of a gene in cell of a subject, by utilizing
PT components of tetracycline repressor/operator inducer system of
PT prokaryotes to regulate gene expression in eukaryotic cells.
XX
PS Disclosure; Fig 7B; 71pp; English.
XX
XX The invention relates to a method of regulating expression of a tet
CC operator-linked gene in cell of a subject, which involves introducing
CC into the cell a nucleic acid molecule encoding a fusion protein which
CC inhibits transcription in eukaryotic cells, comprising a first
CC polypeptide which binds to a tet operator sequence, operatively linked to
CC a heterologous polypeptide which inhibits transcription in eukaryotic
CC cells and modulating concentration of tetracycline, or its analogue in
CC the subject. The method is useful for regulating expression of a
CC exogenous or endogenous gene in a cell. The system has widespread
CC applicability to the study of cellular development and differentiation in
CC eukaryotic cells, plants and animals. For expression of e.g. oncogenes
CC can be regulated in a controlled manner in cells to study their function.

CC The system can be used to regulate the expression of site-specific
CC recombinases such as CRB or Flp, to allow for irreversible modification
CC of the genotype of a transgenic organism under controlled conditions at a
CC particular stage of development. The method is also useful for gene
CC therapy purposes, in treatment for either genetic or acquired diseases,
CC including rheumatoid arthritis, hypopituitarism, wound healing and anti-
CC cancer treatments, for large-scale production of proteins in vitro and in
CC transgenic farm animals and for production of stable cell lines for gene
CC cloning. The present sequence represents DNA encoding a tet promoter
XX

Sequence 520 BP; 152 A; 122 C; 131 G; 115 T; 0 U; 0 Other;

Query Match 5.4%; Score 311.8; DB 8; Length 520;

Best Local Similarity 99.4%; Pred. No. 8.7e-78;

Matches 313; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	4277	GGAACTCGAGTTTACCACTCCCTATCAGTATGAGAAAAAGTGAAGTTCGATTACCAAC	4336
DB	63	GGATCTCGAGTTTACCACTCCCTATCAGTATGAGAAAAAGTGAAGTTCGATTACCAAC	122
QY	4337	TCCTATCAGTATGAGAAAAAGTGAAGTTCGATTACCACTCCCTATCAGTATGAG	4396
DB	123	TCCTATCAGTATGAGAAAAAGTGAAGTTCGATTACCACTCCCTATCAGTATGAG	182
QY	4397	AAAAGTGAAGTTCGAGTTTACCACTCCCTATCAGTATGAGAAAAAGTGAAGTTCGATT	4456
DB	183	AAAAGTGAAGTTCGAGTTTACCACTCCCTATCAGTATGAGAAAAAGTGAAGTTCGATT	242
QY	4457	TACCACTCCCTATCAGTATGAGAAAAAGTGAAGTTCGATTACCACTCCCTATCAGTG	4516
DB	243	TACCACTCCCTATCAGTATGAGAAAAAGTGAAGTTCGATTACCACTCCCTATCAGTG	302
QY	4517	ATGAGAAAAAGTGAAGTTCGAGTTTACCACTCCCTATCAGTATGAGAAAAAGTGAAGT	4576
DB	303	ATGAGAAAAAGTGAAGTTCGAGTTTACCACTCCCTATCAGTATGAGAAAAAGTGAAGT	362
QY	4577	CGAGCTCGGTACCAG	4591
DB	363	CGAGCTCGGTACCCG	377

Search completed: August 29, 2005, 19:54:21
Job time : 1882 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 29, 2005, 17:44:27 ; Search time 15852 Seconds
(without alignments)
17530.322 Million cell updates/sec

Title: US-10-613-728-1
Perfect score: 5735
Sequence: 1 GGATCTGCAAGTCTCACCACACCAAGGTCCTAGTCTCAATTTCACT 60

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBank1:*
1: gb_ba:*
2: gb_hgt:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pac:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sbs:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	5045	88.0	5443	6 BD140880	BD140880 An animal
2	5045	88.0	5443	10 MM071441	U71441 Mus musculu
3	4804.8	83.8	235759	2 AC095977	AC095977 Mus muscu
4	2343	40.9	185702	2 AC130940	AC130940 Rattus no
5	2343	40.9	210784	2 AC119293	AC119293 Rattus no
6	2320.2	40.5	5487	10 AY191158	AY191158 Rattus no
7	1655.4	28.9	263901	2 AC115371	AC115371 Rattus no
8	1444.8	25.2	32415	10 HAMSHCA	L15351 Mesocricetu
9	1279.8	22.3	1679	6 AR054006	AR054006 Sequence
10	1279.8	22.3	1679	6 AR146182	AR146182 Sequence
11	1279.8	22.3	1679	6 AX546496	AX546496 Sequence
12	1279.8	22.3	2594	10 MUSCMHCG	M62404 Mus musculu
13	558.2	9.7	157910	9 CNS01DTR	AL132855 Human chr
14	558.2	9.7	196292	9 CNS0000B	AL049829 Human chr
15	494.8	8.6	31462	9 HSCMCHCA	Z20656 Homo sapien
16	462.6	8.1	516	11 BV093112	BV093112 RPAMMSRO
17	462.6	8.1	542	11 BV163870	BV163870 RPAMMSRO
18	438.4	7.6	452	11 BV093084	BV093084 RPAMMSRO
19	424.4	7.4	641	10 RATMYHAB1	K01463 Rat cardiac

20	407.6	7.1	6020	4 AF192305	AF192305 Oryctolag
C 21	396.4	6.9	541	11 BV159459	BV159459 RPAMMSRO
C 22	395.4	6.9	538	11 BV093111	BV093111 RPAMMSRO
C 23	378	6.6	580	11 BV159460	BV159460 RPAMMSRO
24	377	6.6	581	11 BV093113	BV093113 RPAMMSRO
25	343.6	6.0	2450	6 BD096863	BD096863 Agent for
26	312.6	5.5	10728	6 AR562497	AR562497 Sequence
27	311.8	5.4	520	6 AR020198	AR020198 Sequence
28	311.8	5.4	520	6 AR032150	AR032150 Sequence
29	311.8	5.4	520	6 AR043815	AR043815 Sequence
30	311.8	5.4	520	6 AR072127	AR072127 Sequence
31	311.8	5.4	520	6 AR095981	AR095981 Sequence
32	311.8	5.4	520	6 AR136641	AR136641 Sequence
33	311.8	5.4	520	6 AR157176	AR157176 Sequence
34	311.8	5.4	520	6 AR164125	AR164125 Sequence
35	311.8	5.4	520	6 I12811	I12811 Sequence 7
36	311.8	5.4	520	6 I59625	I59625 Sequence 7
37	311.8	5.4	520	6 AX127248	AX127248 Sequence
38	310.4	5.4	5961	12 AF326320	AF326320 Shuttle v
C 39	310.4	5.4	450	6 AR020200	AR020200 Sequence
C 40	310.4	5.4	450	6 AR029415	AR029415 Sequence
C 41	310.4	5.4	450	6 AR032152	AR032152 Sequence
C 42	310.4	5.4	450	6 AR043817	AR043817 Sequence
C 43	310.4	5.4	450	6 AR072129	AR072129 Sequence
C 44	310.4	5.4	450	6 AR095983	AR095983 Sequence
C 45	310.4	5.4	450	6 AR136643	AR136643 Sequence

ALIGNMENTS

RESULT 1
BD140880
LOCUS BD140880 An animal into which p300 gene is introduced.
DEFINITION BD140880
ACCESSION BD140880.1 GI:23235825
VERSION WO 0205633-A/2.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus

REFERENCE
1 (bases 1 to 5443)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
Hasegawa, K., Kawase, Y. and Suzuki, H.
An animal into which p300 gene is introduced
Patent: WO 0205633-A 2 24-JAN-2002;
KOJI HASEGAWA, CHUGAI PHARMACEUTICAL CO LTD, YOSUKE KAWASE, HIROSHI SUZUKI

COMMENT
OS Mus musculus (mouse)
PN WO 0205633-A/2
PD 24-JAN-2002
PF 13-JUL-2001 WO 2001JP006086
PR 14-JUL-2000 JP 00P 215143
PI KOJI HASEGAWA, YOSUKE KAWASE, HIROSHI SUZUKI
PC A01K67/027, C12N15/09, C12N15/12, C12Q1/68, A61K45/00, A61P9/04, PC

GOIN33/50,
PC GOIN33/15
CC An animal into which p300 gene is introduced
FH Key Location/Qualifiers
FT promoter (1) . (5443).
LOCATION/Qualifiers
1.5443
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"

ORIGIN

Query Match 88.0%; Score 5045; DB 6; Length 5443;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 5363; Conservative 0; Mismatches 60; Indels 29; Gaps 1;
1 GGATCTGCAAGTCTCACCACCAAGGTCCTAGTCTCAATTTCACT 60

BEST AVAILABLE COPY

```

1141 TGAAGACATGGGATTGGAGACATCTCTTTGAGTCTCCCTCAACCCCACTTACAGAC 1200
Db 1141 TGAAGACATGGGATTGGAGACATCTCTTTGAGTCTCCCTCAACCCCACTTACAGAC 1200
Qy 1201 ACACGCTGTGTGGCAGACTCCTGTTCAACAGCCCTGTGTGTCGACCACTGAGCTAG 1260
Db 1201 ACACGCTGTGTGGCAGACTCCTGTTCAACAGCCCTGTGTGTCGACCACTGAGCTAG 1260
Qy 1261 GCAACGAGACATGGGCCCTGTGTGAGATGAGAAGTTGGTTACCAATAGCAAAAACAG 1320
Db 1261 GCAACGAGACATGGGCCCTGTGTGAGATGAGAAGTTGGTTACCAATAGCAAAAACAG 1320
Qy 1321 CAGGGAGGAGAGAACAGAACCAATTAAGAAAGAAAGAAAGCCCACTCAATCAAG 1380
Db 1321 CAGGGAGGAGAGAACAGAACCAATTAAGAAAGAAAGAAAGCCCACTCAATCAAG 1380
Qy 1440 TGTATTTAGAGAGATGGGAGCCACACACAGCTTGAGCAGAGAAACGAAAAGGAG 1440
Db 1440 TGTATTTAGAGAGATGGGAGCCACACACAGCTTGAGCAGAGAAACGAAAAGGAG 1440
Qy 1500 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1500
Db 1500 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1500
Qy 1560 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1560
Db 1560 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1560
Qy 1620 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1620
Db 1620 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1620
Qy 1680 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1680
Db 1680 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1680
Qy 1740 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1740
Db 1740 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1740
Qy 1800 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1800
Db 1800 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1800
Qy 1860 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1860
Db 1860 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1860
Qy 1920 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1920
Db 1920 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1920
Qy 1980 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1980
Db 1980 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 1980
Qy 2040 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 2040
Db 2040 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 2040
Qy 2100 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 2100
Db 2100 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 2100
Qy 2160 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 2160
Db 2160 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 2160
Qy 2220 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 2220
Db 2220 AATTTTATTAAGAGAGCCACAGAAAGAGAGCCCAAGTCTCTCTTT 2220
```


2221 AAGGAAGAGGCGCAACCGGCCACACAGAGGGCAGAGCCAGAACCTGAGTTAATC 2280
2222 AAGGAAGAGGCGCAACCGGCCACACAGAGGGCAGAGCCAGAACCTGAGTTAATC 2280
2281 CTTCCTTGTGCACTTTCATAGAGGAGGAGGAACTCTGTGACCAACATCCCATGA 2340
2282 CTTCCTTGTGCACTTTCATAGAGGAGGAGGAACTCTGTGACCAACATCCCATGA 2340
2341 GCGCCCACTACCCATCCAGATTTGGCTGAGGGAATCTAGTTCCCTGAGAGCAGAG 2400
2342 GCGCCCACTACCCATCCAGATTTGGCTGAGGGAATCTAGTTCCCTGAGAGCAGAG 2400
2401 CCTGGCTTTGTCTCTTGGACCTGACCCCAAGCTGACCCCAATGTTCTCAGTACCTT 2460
2402 CCTGGCTTTGTCTCTTGGACCTGACCCCAAGCTGACCCCAATGTTCTCAGTACCTT 2460
2461 GCGCTCAAGAGCTTGAAGACAGAGCAGTGAATATTTAGGCAATGGGCTAACCCTGAG 2520
2462 GCGCTCAAGAGCTTGAAGACAGAGCAGTGAATATTTAGGCAATGGGCTAACCCTGAG 2520
2521 TGCAACAGAGGCTTGAAGACAGAGCAGTGAATATTTAGGCAATGGGCTTATATCC 2580
2522 TGCAACAGAGGCTTGAAGACAGAGCAGTGAATATTTAGGCAATGGGCTTATATCC 2580
2581 CCAAGAGCAACCAATTTGGCATAGTGGCTGCAATGGGAATGCAAGTTGATCAGTTC 2640
2582 CCAAGAGCAACCAATTTGGCATAGTGGCTGCAATGGGAATGCAAGTTGATCAGTTC 2640
2641 CCTTCAAGAAATCTGCATCAAGACCTAAGACCCCTGAGAGAGGGGTAATGCTCTGCC 2700
2642 CCTTCAAGAAATCTGCATCAAGACCTAAGACCCCTGAGAGAGGGGTAATGCTCTGCC 2700
2701 CCAACCAACCAATTTGGGAATGTAATCTCTTGGGGCTGGCCCTTGGGGAGACACAC 2760
2702 CCAACCAACCAATTTGGGAATGTAATCTCTTGGGGCTGGCCCTTGGGGAGACACAC 2760
2761 ATTACTGAGAGTGTGAGCCAGAAAACCTGACCCGCTGTGTCTGCTGACCACTCCAG 2820
2762 ATTACTGAGAGTGTGAGCCAGAAAACCTGACCCGCTGTGTCTGCTGACCACTCCAG 2820
2821 TCTAGAGCTATATTGAGAGGTGACAGTATAGGGTGGAGCTGTGAGCAGGAGAGTGT 2880
2822 TCTAGAGCTATATTGAGAGGTGACAGTATAGGGTGGAGCTGTGAGCAGGAGAGTGT 2880
2881 TCTGTGGGTGAGGGGTGAGGGGAAAGCCAGAGCAGGGAGTGTGGCTTTGTCTCTGA 2940
2882 TCTGTGGGTGAGGGGTGAGGGGAAAGCCAGAGCAGGGAGTGTGGCTTTGTCTCTGA 2940
2941 CACAATGTCTACTTATTATTAACAGGATGACCTGTAAAGACCCCAACATCTACGACTC 3000
2942 CACAATGTCTACTTATTATTAACAGGATGACCTGTAAAGACCCCAACATCTACGACTC 3000
3001 TGAAGAGACAGCAGCCCTGAGAGCAGGGGTTGTCTGAGCCTTGGGTCTTGTATGTG 3060
3002 TGAAGAGACAGCAGCCCTGAGAGCAGGGGTTGTCTGAGCCTTGGGTCTTGTATGTG 3060
3061 CCAACAAAGAGGAGCATGAGTGTGATTAAGGCCCCAGAGAGCGTTTGAAGAGGCACTTG 3120
3062 CCAACAAAGAGGAGCATGAGTGTGATTAAGGCCCCAGAGAGCGTTTGAAGAGGCACTTG 3120
3121 GGAAGGGGTCAAGTCTGAGAGCCCTATCATGGAATCTGAGCCTTGGGGCCAACTGGTG 3180
3122 GGAAGGGGTCAAGTCTGAGAGCCCTATCATGGAATCTGAGCCTTGGGGCCAACTGGTG 3180
3181 TAAATCTTGTGGCTGCAAGGCAATTCAAAGCAGCACTGATCTCTTGGCAGCTGGGA 3240
3182 TAAATCTTGTGGCTGCAAGGCAATTCAAAGCAGCACTGATCTCTTGGCAGCTGGGA 3240
3241 GGGGGAAGGAGCAACCCCACTTATACCTTTCTCCCTCAGCCCCAGAGATTAAACCTC 3300
3242 GGGGGAAGGAGCAACCCCACTTATACCTTTCTCCCTCAGCCCCAGAGATTAAACCTC 3300
3301 CTGGCTTCCCTTCCACCTCCATCAAGAGTGAAGGGTTGCAAGGAGGGTAAAAA 3360

3301 CTGGCTTCCCTTCCACCTCCATCAGAGTGAAGGGTTGCAAGAGGAGGGTAAAAA 3360
3361 CCTACATGTCCAAACATCATGTGTGACAGATATATGTGATCAGTATGTGAGGCAAGAA 3420
3362 CCTACATGTCCAAACATCATGTGTGACAGATATATGTGATCAGTATGTGAGGCAAGAA 3420
3421 GGAATCTGAGGCTTAATGGGTTAATGTGTAAGGTCGTGTGATGATGTGTGTCTG 3480
3422 GGAATCTGAGGCTTAATGGGTTAATGTGTAAGGTCGTGTGATGATGTGTGTCTG 3480
3481 ACTGAAACCGGCAAGCTGTGACAGCTGTGATGATGATGATGATGATGATGATGATG 3540
3482 ACTGAAACCGGCAAGCTGTGACAGCTGTGATGATGATGATGATGATGATGATGATG 3540
3541 GGTGTGTGTAAATTTGCCCAAGGCAAGTGGTGAATCCCTTCAATGTTTAAAGAT 3600
3542 GGTGTGTGTAAATTTGCCCAAGGCAAGTGGTGAATCCCTTCAATGTTTAAAGAT 3600
3601 TGGATGATGGCTGATCTCAAGAGCCATGGAAATGAAATGGAACCTATATGTGTCT 3660
3602 TGGATGATGGCTGATCTCAAGAGCCATGGAAATGAAATGGAACCTATATGTGTCT 3660
3661 CTAACTAAGGTAGCAAGGCTTTTGGAGAGACCTGTCTAAGATGTGGCAACAGAGAC 3720
3662 CTAACTAAGGTAGCAAGGCTTTTGGAGAGACCTGTCTAAGATGTGGCAACAGAGAC 3720
3721 TACAGACATGTCTGTACAGATTAAGAGAGAGAGAGAGGGGTGTAAATTTCTTTACTA 3780
3722 TACAGACATGTCTGTACAGATTAAGAGAGAGAGAGAGGGGTGTAAATTTCTTTACTA 3780
3781 TCAAGAGAACTGAGTGTGACCTGCAAGTGAATGCTCTCCCTAGACATCAGACTT 3840
3782 TCAAGAGAACTGAGTGTGACCTGCAAGTGAATGCTCTCCCTAGACATCAGACTT 3840
3841 TGTCTCTGGGAGCCAGACCTGTGAACTTCAAGTCTGAGAGAGTGAAGGCTCTCCCTCA 3900
3842 TGTCTCTGGGAGCCAGACCTGTGAACTTCAAGTCTGAGAGAGTGAAGGCTCTCCCTCA 3900
3901 GCTGAAGCTATGACAGATTAAGGCTTGAAGGAGGAGAGGAGGCTGGATGGAGAC 3960
3902 GCTGAAGCTATGACAGATTAAGGCTTGAAGGAGGAGAGGAGGCTGGATGGAGAC 3960
3961 TGTGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4020
3962 TGTGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4020
4021 TTGTTCAACTCACTTCAAGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4080
4022 TTGTTCAACTCACTTCAAGATTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4080
4081 GTGTGAGAGCTCTGTGTCTCTCTGTGATGCTGAGGCTTGGGAGGAGGAGTGT 4140
4082 GTGTGAGAGCTCTGTGTCTCTCTGTGATGCTGAGGCTTGGGAGGAGGAGTGT 4140
4083 GTGTGAGAGCTCTGTGTCTCTCTGTGATGCTGAGGCTTGGGAGGAGGAGTGT 4140
4141 GCGCAAGAGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4200
4142 GCGCAAGAGCTTAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 4200
4201 GCAAACTTGGGGCCCTGTGCTCTCTGTGATGCTGAGGCTTGGGAGGAGGAGTGT 4260
4202 GCAAACTTGGGGCCCTGTGCTCTCTGTGATGCTGAGGCTTGGGAGGAGGAGTGT 4260
4261 TTCTAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4320
4262 TTCTAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4320
4321 AAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAGTGAAGTCCAGTTTACCACTC 4380
4322 AAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAGTGAAGTCCAGTTTACCACTC 4380
4381 CCTATCAGTATAGAGAAAGTGAAGTCCAGTTTACCACTCCCTATCAGTATAGAGAA 4440

```

Db 4292 ----- 4291
Qy 4441 AAGTGAAGTGAAGTTACCACTCCATCAGTATAGAGAAAAGTGAAGTGAATTGA 4500
Db 4292 ----- 4291
Qy 4501 CCATCCCTATCAGTATAGAGAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4560
Db 4292 ----- 4291
Qy 4561 AGAGAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4620
Db 4292 ----- 4291
Qy 4621 TATGGATGGATATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4680
Db 4329 TATGGATGGATATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4388
Qy 4681 GGTAAAGGAGATTTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4740
Db 4389 GGTAAAGGAGATTTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4448
Qy 4741 GGAAGAGAGAGATTTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4800
Db 4449 GGAAGAGAGAGATTTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4508
Qy 4801 CCAACAGAGAGATTTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4860
Db 4509 CCAACAGAGAGATTTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 4568
Qy 4861 TCCCTTACGCTGCTTCTCCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4920
Db 4569 TCCCTTACGCTGCTTCTCCATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4628
Qy 4921 TTGGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4980
Db 4629 TTGGCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4688
Qy 4981 CTCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5040
Db 4689 CTCGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4748
Qy 5041 TTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5100
Db 4749 TTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4808
Qy 5101 CTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5160
Db 4809 CTTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4868
Qy 5161 CTGTGTGAGAGTGTGAGATCAGACCTGGGGTTCACACCTTATGTAAACAATCTTCCA 5220
Db 4869 CTGTGTGAGAGTGTGAGATCAGACCTGGGGTTCACACCTTATGTAAACAATCTTCCA 4928
Qy 5221 GTGAGCCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 5280
Db 4929 GTGAGCCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 4988
Qy 5281 CTGTTCATCTGTGTGAGATCTCTAGATTGTGTCTCCAGAGCTTGTGTAATCTCTCTCTCTCT 5340
Db 4989 CTGTTCATCTGTGTGAGATCTCTAGATTGTGTCTCCAGAGCTTGTGTAATCTCTCTCTCTCTCT 5048
Qy 5341 GCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5400
Db 5049 GCGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5108
Qy 5401 ATTCTTCAAGATTTCTCTGAAAAGTTAAACAGAGTGAAGATTTTCTCTCTCTCTCTCTCTCTCT 5460
Db 5109 ATTCTTCAAGATTTCTCTGAAAAGTTAAACAGAGTGAAGATTTTCTCTCTCTCTCTCTCTCTCT 5168
Qy 5461 TCACGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5520
Db 5169 TCACGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5228

```

```

Qy 5521 CTCTTAGCAAAAGCTCAGGACACCTTTACCCCAATAGACCTCTGACAGAGAAGGAGCACT 5580
Db 5229 CTCTTAGCAAAAGCTCAGGACACCTTTACCCCAATAGACCTCTGACAGAGAAGGAGCACT 5288
Qy 5581 TTACATGAGTCTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5640
Db 5289 TTACATGAGTCTCTGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5348
Qy 5641 GGTGTAGGAAAGTCAAGAGCTTTCATAGAAAGCTTACAGCCACAGAAATGACAGACAG 5700
Db 5349 GGTGTAGGAAAGTCAAGAGCTTTCATAGAAAGCTTACAGCCACAGAAATGACAGACAG 5408
Qy 5701 TCCCTCTATCTCTCCCATTAAGAGTTTGAAGTGCAC 5735
Db 5409 TCCCTCTATCTCTCCCATTAAGAGTTTGAAGTGCAC 5443

```

```

RESULT 2
MMU71441
LOCUS MMU71441 5443 bp DNA linear ROD 21-OCT-1996
DEFINITION Mus musculus alpha myosin heavy chain gene, promoter region.
ACCESSION U71441
VERSION U71441.1 GI:1621436
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)
REFERENCE
AUTHORS Gulick,J. and Robbins,J.
TITLE Mouse myosin heavy chain promoter
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 5443)
AUTHORS Gulick,J. and Robbins,J.
TITLE Direct Submision
JOURNAL Submitted (20-SEP-1996) Molecular Card. Biol., Children's Hospital, 3333 Burnet Avenue, Cincinnati, OH 45229-3039, USA
FEATURES
source
1..5443
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
1..5443
/note="alpha myosin heavy chain; capable of driving high levels of cardiac specific transcription in vitro or in vivo"

```

promoter

ORIGIN

Query Match 88.0%; Score 5045; DB 10; Length 5443;
 Best Local Similarity 93.9%; Pred. No. 0;
 Matches 5383; Conservative 0; Mismatches 60; Indels 292; Gaps 1;

```

Qy 1 GGAATCCGCAAGGTCAACAAGGCTCTCCACCCAGAGAGGCTGTGATCAATTTCACT 60
Db 1 GGAATCCGCAAGGTCAACAAGGCTCTCCACCCAGAGAGGCTGTGATCAATTTCACT 60
Qy 61 TTCCATGCTTGTCTCAACATCTGGGCTCCCAAGCTAATTTGAGCTTTGTTTAT 120
Db 61 TTCCATGCTTGTCTCAACATCTGGGCTCCCAAGCTAATTTGAGCTTTGTTTAT 120
Qy 121 TTCAAAAGGCTGAAATGAGAGTATGTTGTCTACCCAGCTCTTAAGGATGCCGTGA 180
Db 121 TTCAAAAGGCTGAAATGAGAGTATGTTGTCTACCCAGCTCTTAAGGATGCCGTGA 180
Qy 181 AGCCCTCAACCTGGAGCCCTTTCACAACAGCCCTTTAGGTGGAAGAGAAATTAAGCAATT 240
Db 181 AGCCCTCAACCTGGAGCCCTTTCACAACAGCCCTTTAGGTGGAAGAGAAATTAAGCAATT 240
Qy 241 TCCTTAAGCAAAATCTGCTCTGAGCTCTTCTCTGACCTGCTGCTCCCTGAGCTCT 300
Db 241 TCCTTAAGCAAAATCTGCTCTGAGCTCTTCTCTGACCTGCTGCTCCCTGAGCTCT 300

```

QY 301 AGGTTGGGAGTGTGGGCTTGAGAGAAAGTGGGAAAGTGGCAAAAGCCATCCCTAG 360
DB 301 AGGTTGGGAGTGTGGGCTTGAGAGAAAGTGGGAAAGTGGCAAAAGCCATCCCTAG 360
QY 361 GGCCTGTGAGTGTGGAGCTTTCCTGTACAGACTGGCTCATAGATCTCTCCAGCC 420
DB 361 GGCCTGTGAGTGTGGAGCTTTCCTGTACAGACTGGCTCATAGATCTCTCCAGCC 420
QY 421 AAACATAGCAAGAGTATCTCTTGTGATTTCCCGAGGCCAGTACCTGTCAAGTT 480
DB 421 AAACATAGCAAGAGTATCTCTTGTGATTTCCCGAGGCCAGTACCTGTCAAGTT 480
QY 481 GAAACAGATTTTGAAGAGCTCTGAACCTGAACCTGAAGCTCATCCACAGCA 540
DB 481 GAAACAGATTTTGAAGAGCTCTGAACCTGAACCTGAAGCTCATCCACAGCA 540
QY 541 AGCAGCTAGGTGCACCTGCTAGTTAGTATCTACGCTGATATATGCAAGCTGGGCCAC 600
DB 541 AGCAGCTAGGTGCACCTGCTAGTTAGTATCTACGCTGATATATGCAAGCTGGGCCAC 600
QY 601 AGAAGTCTGGGGGTGAGAACTGACCAAGTACTTTTCACTGGGCAAAAGGTATGACCCC 660
DB 601 AGAAGTCTGGGGGTGAGAACTGACCAAGTACTTTTCACTGGGCAAAAGGTATGACCCC 660
QY 661 TCAGAGATGTAGTATGTCCCTTAGATCCATCCAGGAGAGTCTTAAGAGACATG 720
DB 661 TCAGAGATGTAGTATGTCCCTTAGATCCATCCAGGAGAGTCTTAAGAGACATG 720
QY 721 GGATGAGAGTGTAGTATGTGTCATTCACACAGCTATCCACAGTGTCCCTTGGCCC 780
DB 721 GGATGAGAGTGTAGTATGTGTCATTCACACAGCTATCCACAGTGTCCCTTGGCCC 780
QY 781 TTCCACTTGGCCAGAGAGCAAGTAACTTATCTTCTTCCCTCCATCTCCAG 840
DB 781 TTCCACTTGGCCAGAGAGCAAGTAACTTATCTTCTTCCCTCCATCTCCAG 840
QY 841 GACACACCCCTGTGTGAGTATTCATTTCTTCACTGCTCCCTGTGTGACTTCAT 900
DB 841 GACACACCCCTGTGTGAGTATTCATTTCTTCACTGCTCCCTGTGTGACTTCAT 900
QY 901 TTGCAAGGCTTTTTCAGCTGTGAGCTGTGGAAGATAGAGTTTGGCTTAGGTGTGCA 960
DB 901 TTGCAAGGCTTTTTCAGCTGTGAGCTGTGGAAGATAGAGTTTGGCTTAGGTGTGCA 960
QY 961 GGCATCTCAAGAGAAAGCAACAGAGGGGACCAATTTTGAAGATAGAGAACTTAA 1020
DB 961 GGCATCTCAAGAGAAAGCAACAGAGGGGACCAATTTTGAAGATAGAGAACTTAA 1020
QY 1021 TCAGTGGCGGGCTGGGGGTGAGAAAGAGTGAAGTCCGCTCAGCTAGCCAGC 1080
DB 1021 TCAGTGGCGGGCTGGGGGTGAGAAAGAGTGAAGTCCGCTCAGCTAGCCAGC 1080
QY 1081 TAGTCCCGAGATCTCTGACAGCTGGCTGTGGGGGTAGCTTTAGGAATGTGGGTC 1140
DB 1081 TAGTCCCGAGATCTCTGACAGCTGGCTGTGGGGGTAGCTTTAGGAATGTGGGTC 1140
QY 1141 TGAAGACATTTGGATTTGGAAGCATCTTTTGAAGTCTCCCTCAACCCCACTTACAG 1200
DB 1141 TGAAGACATTTGGATTTGGAAGCATCTTTTGAAGTCTCCCTCAACCCCACTTACAG 1200
QY 1201 ACAGCTGTGTGGCGAGCTCTGCTTCAACAGCCCTCTGTGTCTTGAACACTGAGCTAG 1260
DB 1201 ACAGCTGTGTGGCGAGCTCTGCTTCAACAGCCCTCTGTGTCTTGAACACTGAGCTAG 1260
QY 1261 GCAACAGAGCATGGGCTGTGTCTGAGATGAAGAGTTGTATCAAAATGCAAAAACAG 1320
DB 1261 GCAACAGAGCATGGGCTGTGTCTGAGATGAAGAGTTGTATCAAAATGCAAAAACAG 1320
QY 1321 CAGGGAGGGAGAAACAGAAACAAATTAAGAGAGAAAGGAGCCAGTCAATCAG 1380
DB 1321 CAGGGAGGGAGAAACAGAAACAAATTAAGAGAGAAAGGAGCCAGTCAATCAG 1380
QY 1381 TGCAGTCAAGAGATGGAGGCCAACACAGCTTGAAGAGAGAAACGAAAAGGGAG 1440

DB 1381 TGCAGTCAAGAGATGGAGGCCAACACAGCTTGAAGAGAGAAACGAAAAGGGAG 1440
QY 1441 AGATTCTGGGCTTAAGAGAGCCACAGAAAGAGCCAGGCCCTCCAGTCTCTCTTT 1500
DB 1441 AGATTCTGGGCTTAAGAGAGCCACAGAAAGAGCCAGGCCCTCCAGTCTCTCTTT 1500
QY 1501 ATACCCTATCCCGCTCCCAATTAAAGCCACTCTTCTTCTAGATCAGACTGAGCTGC 1560
DB 1501 ATACCCTATCCCGCTCCCAATTAAAGCCACTCTTCTTCTAGATCAGACTGAGCTGC 1560
QY 1561 AGCGAAGAGACCCGTAAGAGAGATCACTGATAGAGAGATGTGTGAGAGATCCAG 1620
DB 1561 AGCGAAGAGACCCGTAAGAGAGATCACTGATAGAGAGATGTGTGAGAGATCCAG 1620
QY 1621 GCAACCTTAAGAGCCAGAGCTTAAAGAGCAGAGATTAAGGTCTTCAAGGTGGCCAG 1680
DB 1621 GCAACCTTAAGAGCCAGAGCTTAAAGAGCAGAGATTAAGGTCTTCAAGGTGGCCAG 1680
QY 1681 CTGTGACACAGAGGCTGAGAGCTGTGTGAGCTTCAAGATTAAGATGATGCTCAGA 1740
DB 1681 CTGTGACACAGAGGCTGAGAGCTGTGTGAGCTTCAAGATTAAGATGATGCTCAGA 1740
QY 1741 ATGGCGGGGGGGGATTTCTGGGGGGGGAGAGAAAGTGTGAGAAAGCCTGAGACAG 1800
DB 1741 ATGGCGGGGGGGGATTTCTGGGGGGGGAGAGAAAGTGTGAGAAAGCCTGAGACAG 1800
QY 1801 AGAATCTGAGAGCCTGGAACGATTAACATTAAGAGAGAACCCAGGCTTATGATG 1860
DB 1801 AGAATCTGAGAGCCTGGAACGATTAACATTAAGAGAGAACCCAGGCTTATGATG 1860
QY 1861 TAAATCATGAAGACAGAGGAGAGAGGAAAGCTGGAAGATGAAGAACCCCGGGCAGA 1920
DB 1861 TAAATCATGAAGACAGAGGAGAGAGGAAAGCTGGAAGATGAAGAACCCCGGGCAGA 1920
QY 1921 CATGAGAGCAAGCAAGCCAGGTTGAGGCTCCGTGAATCAGCTGTGAGAGCAGAG 1980
DB 1921 CATGAGAGCAAGCAAGCCAGGTTGAGGCTCCGTGAATCAGCTGTGAGAGCAGAG 1980
QY 1981 CCTGGTATGAGCACCAAGACGAGAGGCTTAGGCTTAATGTGAGACAGAGAAACAG 2040
DB 1981 CCTGGTATGAGCACCAAGACGAGAGGCTTAGGCTTAATGTGAGACAGAGAAACAG 2040
QY 2041 GTAGACACAGAAACAGACAGAGAGCGGGGAGCCAGGTAAACAAAGAAATGCTCTGAC 2100
DB 2041 GTAGACACAGAAACAGACAGAGAGCGGGGAGCCAGGTAAACAAAGAAATGCTCTGAC 2100
QY 2101 CTGTGGCAGAGCTGCATCTGTGTCAATATCTTAGAATGTTCAATCAGACTGACAGGC 2160
DB 2101 CTGTGGCAGAGCTGCATCTGTGTCAATATCTTAGAATGTTCAATCAGACTGACAGGC 2160
QY 2161 TGGCTTGGAGGACCTGGAAGAGATGTGAGAGCCAGGGGAGCAAGGGGGCTTAGGA 2220
DB 2161 TGGCTTGGAGGACCTGGAAGAGATGTGAGAGCCAGGGGAGCAAGGGGGCTTAGGA 2220
QY 2221 AAGGAAGAGAGGGGAAACAGGCCACACAAAGGGGAGAGGCCAGAACTGAGTTAACTC 2280
DB 2221 AAGGAAGAGAGGGGAAACAGGCCACACAAAGGGGAGAGGCCAGAACTGAGTTAACTC 2280
QY 2281 CTTCCTTGTGATCTTCCATAGAGAGGAGTGGAACTGTGACACCATCCCCATGA 2340
DB 2281 CTTCCTTGTGATCTTCCATAGAGAGGAGTGGAACTGTGACACCATCCCCATGA 2340
QY 2341 GCGCCCACTAACCATACCAAGTTTGGCTGAGTGGCATTTAGTTCTCTGAGACAGAG 2400
DB 2341 GCGCCCACTAACCATACCAAGTTTGGCTGAGTGGCATTTAGTTCTCTGAGACAGAG 2400
QY 2401 CCTGGCCTTTGTCTCTGAGACCTGACCCAAAGCTGACCCCAATTTCTCACTTATCAT 2460
DB 2401 CCTGGCCTTTGTCTCTGAGACCTGACCCAAAGCTGACCCCAATTTCTCACTTATCAT 2460
QY 2461 GCGCTCAAGAGCTTAGAACCAAGAGTGAATATTAAGCCATGAGCTTACCCTGAGCT 2520

Db 2461 GCCCTCAAGAGCTTGAGAACGAGCAGTACATATTAGGCCAGGCGCTAACCTCGAGCT 2520
Qy 2521 TGCACACAGAGAGCTTCAAGTACCTCCAGGGACACAGCTGCACAGAGTGGCTTTATCC 2580
Db 2521 TGCACACAGAGAGCTTCAAGTACCTCCAGGGACACAGCTGCACAGAGTGGCTTTATCC 2580
Qy 2581 CCAAGAGCAACCATTTGGCATTAGTGGCTGCACATGGGAATGCAAGGTTGAATCAGGTC 2640
Db 2581 CCAAGAGCAACCATTTGGCATTAGTGGCTGCACATGGGAATGCAAGGTTGAATCAGGTC 2640
Qy 2641 CCTTCAAGAAATACGTCATGCAAGCCTTAAGACCCCTGGAGAGGGGTATGCTCTGCCC 2700
Db 2641 CCTTCAAGAAATACGTCATGCAAGCCTTAAGACCCCTGGAGAGGGGTATGCTCTGCCC 2700
Qy 2701 CCAACCAACCATTAAGGGAGTGAATCTATCTAGGGGGCTGGCGACTTGGGGAGACACAC 2760
Db 2701 CCAACCAACCATTAAGGGAGTGAATCTATCTAGGGGGCTGGCGACTTGGGGAGACACAC 2760
Qy 2761 ATTACTGAGAGTGTGAGGCCCGAAGAACTGACCGCCCTGTGTCTGCCACCTCCACAC 2820
Db 2761 ATTACTGAGAGTGTGAGGCCCGAAGAACTGACCGCCCTGTGTCTGCCACCTCCACAC 2820
Qy 2821 TCTAGAGCTATATTGAGAGGTGACAGTAGATAGGGTGGAGCTGGTAGCAGGAGAGTGT 2880
Db 2821 TCTAGAGCTATATTGAGAGGTGACAGTAGATAGGGTGGAGCTGGTAGCAGGAGAGTGT 2880
Qy 2881 TCCTGGGTGTGAGGGGTGAGGGGAAGCCAGAGCAGGGGAGTCTGGCTTTGTCTCTGAA 2940
Db 2881 TCCTGGGTGTGAGGGGTGAGGGGAAGCCAGAGCAGGGGAGTCTGGCTTTGTCTCTGAA 2940
Qy 2941 CACAACTCTACTTAAATTATTAACAGGAGTACCTGTCTAAGAACCAACATCTACGACTC 3000
Db 2941 CACAACTCTACTTAAATTATTAACAGGAGTACCTGTCTAAGAACCAACATCTACGACTC 3000
Qy 3001 TGAAGAGCAGCAGCCCTGAGAGCAGGGGTGTCTCTGAGCCTTGGGGTGTGATGATG 3060
Db 3001 TGAAGAGCAGCAGCCCTGAGAGCAGGGGTGTCTCTGAGCCTTGGGGTGTGATGATG 3060
Qy 3061 CCACAAAGAGAGGAGATGATGTAGTAAAGCCCAAGAGCGTTAAGAAAGGCACTTG 3120
Db 3061 CCACAAAGAGAGGAGATGATGTAGTAAAGCCCAAGAGCGTTAAGAAAGGCACTTG 3120
Qy 3121 GGAAGGGGTCACTGTGCAAGAGCCCTTATCCATGGAATCTGAGCCTGGGGCAACTG 3180
Db 3121 GGAAGGGGTCACTGTGCAAGAGCCCTTATCCATGGAATCTGAGCCTGGGGCAACTG 3180
Qy 3181 TAAATCTCTGGGCGCTGCAAGGCACTTCAAGAGCAGCCTGCTGGCAGCCTGGGGA 3240
Db 3181 TAAATCTCTGGGCGCTGCAAGGCACTTCAAGAGCAGCCTGCTGGCAGCCTGGGGA 3240
Qy 3241 GGCAGGAAGAGCAACCCCACTTATACCTTTCTCCCTCAGCCCGAGATTAACACT 3300
Db 3241 GGCAGGAAGAGCAACCCCACTTATACCTTTCTCCCTCAGCCCGAGATTAACACT 3300
Qy 3301 CTGGGCTTTCCCTTTCCCACTCCATCAGAGATGAGAGGTTGCAAGGGGTAAAA 3360
Db 3301 CTGGGCTTTCCCTTTCCCACTCCATCAGAGATGAGAGGTTGCAAGGGGTAAAA 3360
Qy 3361 CCTACATGTCACAAATCATGATGATGATGATGATGATGATGATGATGATGATGATG 3420
Db 3361 CCTACATGTCACAAATCATGATGATGATGATGATGATGATGATGATGATGATGATG 3420
Qy 3421 GGAATCTGAGAGGTTAACTGGGTTAAATGTGTAAGTCTGTGTCATGTGTGTCTG 3480
Db 3421 GGAATCTGAGAGGTTAACTGGGTTAAATGTGTAAGTCTGTGTCATGTGTGTCTG 3480
Qy 3481 ACTGAAAAAGGGGATGAGTGTGAGCTGTTCAGTTCTGTGTGTGTGTGTGTGTGTGT 3540
Db 3481 ACTGAAAAAGGGGATGAGTGTGAGCTGTTCAGTTCTGTGTGTGTGTGTGTGTGTGT 3540
Qy 3541 GGTGTGTGTGTAAATGCGCAAGGCAAAAGGGTGAATCCCTTGCATGTGTTAAAGAT 3600
Db 3541 GGTGTGTGTGTAAATGCGCAAGGCAAAAGGGTGAATCCCTTGCATGTGTTAAAGAT 3600

Qy 3601 TGGATATGGCCCTGCATCTCAAGACCATGAAATAGATGACACTTATATGTGTCT 3660
Db 3601 TGGATATGGCCCTGCATCTCAAGACCATGAAATAGATGACACTTATATGTGTCT 3660
Qy 3661 CTAAAGTAAAGTAAAGTCTTTTGAAGACACCTGTCTAAGATGTGGCAACAGAGC 3720
Db 3661 CTAAAGTAAAGTAAAGTCTTTTGAAGACACCTGTCTAAGATGTGGCAACAGAGC 3720
Qy 3721 TACAGACAGTATCTGACAGATTAAGAGAGAGAGGGGGGTGTAATTTCTTACTA 3780
Db 3721 TACAGACAGTATCTGACAGATTAAGAGAGAGAGGGGGGTGTAATTTCTTACTA 3780
Qy 3781 TCAAGGAAAAGTGAAGTGTGACCTGCAAAAGTGAATGCTTCTTACATCATGACTT 3840
Db 3781 TCAAGGAAAAGTGAAGTGTGACCTGCAAAAGTGAATGCTTCTTACATCATGACTT 3840
Qy 3841 TGTCTCTGGGAGACCACTGTGAACTTCAAGTCTGAGAGTGAAGAGGCTCCCTCA 3900
Db 3841 TGTCTCTGGGAGACCACTGTGAACTTCAAGTCTGAGAGTGAAGAGGCTCCCTCA 3900
Qy 3901 GCCTGAAGCTATGACAGATAGCCAGGGTTGAAAAGGGGAAAGGAGCCTGGATGGAGC 3960
Db 3901 GCCTGAAGCTATGACAGATAGCCAGGGTTGAAAAGGGGAAAGGAGCCTGGATGGAGC 3960
Qy 3961 TTGTGTGTGAGAGCAGGGACAGATATTAAAGCTGGAAGAGAGTGAACCTTACCA 4020
Db 3961 TTGTGTGTGAGAGCAGGGACAGATATTAAAGCTGGAAGAGAGTGAACCTTACCA 4020
Qy 4021 TTGTTCATCAACCTTCAAGTAAATTAATCACTGAGAGTAAAGGCTGGGTGAGGGAG 4080
Db 4021 TTGTTCATCAACCTTCAAGTAAATTAATCACTGAGAGTAAAGGCTGGGTGAGGGAG 4080
Qy 4081 GTGTGAGAGCCTGCTGTCTCTCTGATGAGCCTGAGGCTTTGGGAGAGAGATGT 4140
Db 4081 GTGTGAGAGCCTGCTGTCTCTCTGATGAGCCTGAGGCTTTGGGAGAGAGATGT 4140
Qy 4141 GCCCAAGACCTAAAAAAGGCGATGAGCCAGAGGGGCGAGGCAACAGACTTTTATG 4200
Db 4141 GCCCAAGACCTAAAAAAGGCGATGAGCCAGAGGGGCGAGGCAACAGACTTTTATG 4200
Qy 4201 GCAAACTTGGGGCCCTGCTGTCTCTCTGATGAGCCTGAGGCTTTGGGAGAGAG 4260
Db 4201 GCAAACTTGGGGCCCTGCTGTCTCTCTGATGAGCCTGAGGCTTTGGGAGAGAG 4260
Qy 4261 TTCTAAGGACAGAGAGGAACTGAGTTTACACTCCTATCAGTATAGAGAAAAGTGA 4320
Db 4261 GAGCCAGACAGAGAGGAAAGTGGAGGAG----- 4320
Qy 4321 AAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTTTACCACTC 4380
Db 4321 AAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTTTACCACTC 4380
Qy 4381 CCTATCAGTATAGAGAAAAGTGAAGTTTACCACTCCCTATCAGTATAGAGAA 4440
Db 4381 CCTATCAGTATAGAGAAAAGTGAAGTTTACCACTCCCTATCAGTATAGAGAA 4440
Qy 4441 AAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTTTAA 4500
Db 4441 AAGTGAAGTGAAGTTTACCACTCCCTATCAGTATAGAGAAAAGTGAAGTTTAA 4500
Qy 4501 CCACTCCCTATCAGTATAGAGAAAAGTGAAGTTTACCACTCCCTATCAGTAT 4560
Db 4501 CCACTCCCTATCAGTATAGAGAAAAGTGAAGTTTACCACTCCCTATCAGTAT 4560
Qy 4561 AGAGAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4620
Db 4561 AGAGAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4620
Qy 4621 TATGGATGGATATTAAGGGGCTGAGACACTGAGAGCTGTCAAGATTTCTCCAACTCA 4680
Db 4621 TATGGATGGATATTAAGGGGCTGAGACACTGAGAGCTGTCAAGATTTCTCCAACTCA 4680
Qy 4680 TATGGATGGATATTAAGGGGCTGAGACACTGAGAGCTGTCAAGATTTCTCCAACTCA 4740
Db 4680 TATGGATGGATATTAAGGGGCTGAGACACTGAGAGCTGTCAAGATTTCTCCAACTCA 4740

QY 4681 GGTAAAGAGAGTTGGGAGTGGGGGCTCTTCAACCCACACAGAGCTCTCCCACTAGAA 4740
 DB 4389 GGTAAAGAGAGTTGGGAGTGGGGGCTCTTCAACCCACACAGAGCTCTCCCACTAGAA 4448
 QY 4741 GGAAGCTGCTTTCTGGAAGTGGGGGCTTCAAGCCGGTCAAGATCTGACAGGGTGCCTT 4800
 DB 4449 GGAAGCTGCTTTCTGGAAGTGGGGGCTTCAAGCCGGTCAAGATCTGACAGGGTGCCTT 4508
 QY 4801 CCACAGCCCTGGAGAGTCTCACTGACAGAGGTTTCCACAAAGAACACTGATGAGCCCT 4860
 DB 4509 CCACAGCCCTGGAGAGTCTCACTGACAGAGGTTTCCACAAAGAACACTGATGAGCCCT 4568
 QY 4861 TCCCTTACGCTCTCTCCATCTTCTCTGAGAGTCTCTCCCGCTTGTGTTATC 4920
 DB 4569 TCCCTTACGCTCTCTCCATCTTCTCTGAGAGTCTCTCCCGCTTGTGTTATC 4628
 QY 4921 TTGGCTCTTCTGCTTCCAGCAAGATTTGGCTGTGCTGCTCACTCACTTCTCACTGT 4980
 DB 4629 TTGGCTCTTCTGCTTCCAGCAAGATTTGGCTGTGCTGCTCACTCACTTCTCACTGT 4688
 QY 4981 CTCCGAGCTTGGCTTGGCTTGGTGTCTCTTCCAGCCATTTCTCACTCAC 5040
 DB 4689 CTCCGAGCTTGGCTTGGCTTGGTGTCTCTTCCAGCCATTTCTCACTCAC 4748
 QY 5041 TTTTCTCCCTCTCTCAATTTGATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5100
 DB 4749 TTTTCTCCCTCTCTCAATTTGATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4808
 QY 5101 TTTTCTCCCT 5160
 DB 4809 TTTTCTCCCT 4868
 QY 5161 CTCTGTCAGAGTCTGAGATTCACACCTGGGGTCTCCACCTGATTTGAAACATCTTCA 5220
 DB 4869 CTCTGTCAGAGTCTGAGATTCACACCTGGGGTCTCCACCTGATTTGAAACATCTTCA 4928
 QY 5221 GTGAGCACAAGTCTGAGTCTGAGTCTCTTCACTTCTCTCACTCTCTCTCTCTCTCT 5280
 DB 4929 GTGAGCACAAGTCTGAGTCTGAGTCTCTTCACTTCTCTCACTCTCTCTCTCTCTCT 4988
 QY 5281 CTGTTTCATCTGTCAGATCTTCAATTTGATTTCTCCAGCTCTGCTACTCTCTCTCT 5340
 DB 4989 CTGTTTCATCTGTCAGATCTTCAATTTGATTTCTCCAGCTCTGCTACTCTCTCTCTCT 5048
 QY 5341 GCTGTCT 5400
 DB 5049 GCTGTCT 5108
 QY 5401 ATTCTTCAGATTTCTGAAAGTTTAAACAGAGTGAAGATTTTCCCTGTAGACAGAGA 5460
 DB 5109 ATTCTTCAGATTTCTGAAAGTTTAAACAGAGTGAAGATTTTCCCTGTAGACAGAGA 5168
 QY 5461 TCAAGATTTCTCCGGAAGTCAAGGCTTCCAGCCCTCTCTCTCTCTCTCTCTCTCTCT 5520
 DB 5169 TCAAGATTTCTCCGGAAGTCAAGGCTTCCAGCCCTCTCTCTCTCTCTCTCTCTCTCT 5228
 QY 5521 CTCTTACCAAACTCTGAGGACCTCTTACCCACATATGACTCTGACAGAGAGGACACT 5580
 DB 5229 CTCTTACCAAACTCTGAGGACCTCTTACCCACATATGACTCTGACAGAGAGGACACT 5288
 QY 5581 TTACATGAGATCTGTCGAGAGAGCCTTACGCTTAAAGGAGAGGAGGAGTGTGT 5640
 DB 5289 TTACATGAGATCTGTCGAGAGAGCCTTACGCTTAAAGGAGAGGAGGAGTGTGT 5348
 QY 5641 GGTGTAGAAAGTCAAGAGCTTCAATGAAGCTTACGCCACACAGAAATGACAGACAGA 5700
 DB 5349 GGTGTAGAAAGTCAAGAGCTTCAATGAAGCTTACGCCACACAGAAATGACAGACAGA 5408
 QY 5701 TCCCTCTATCTCCCACTAAGATTTGATTCAC 5735
 DB 5409 TCCCTCTATCTCCCACTAAGATTTGATTCAC 5443

RESULT 3
 AC099577
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS

AC099577 235759 bp DNA linear HTG 22-FEB-2004
 Mus musculus chromosome 14 clone Rp23-171A13 map 14, *** SEQUENCING
 IN PROGRESS ***, 6 unordered pieces.
 AC099577
 AC099577.3 GI:42734575
 HTG PHASE1; HTGS FULLTOP; HTGS ACTIVERIN.
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 235759)
 Birren, B., Nusbaum, C. and Lander, E.
 Mus musculus chromosome 14, clone Rp23-171A13
 Unpublished
 2 (bases 1 to 235759)
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B.,
 Brown, A., Camarata, J., Campilano, A., Chang, J., Chazaro, B.,
 Choquel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A.,
 Cooke, P., DeRellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,
 Ginde, S., Gird, S., Goyette, M., Graham, L., Grand-Pierre, N.,
 Hages, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
 Jones, C., Kamat, A., Karatas, A., Kelle, C., LaRocque, K.,
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
 Maclean, C., MacDonald, P., Major, J., Margis, N., Matthews, C.,
 McCarthy, M., McEwan, P., McKernan, K., McPherson, R., Meldrum, J.,
 Menes, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
 Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
 Raymond, C., Retta, M., Rieback, M., Riley, R., Rise, C., Rogov, P.,
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnepf, R.,
 Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
 Strauss, N., Subramanian, A., Talamas, J., Testafaye, S., Theodore, J.,
 Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
 Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (16-NOV-2001) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 235759)
 Birren, B., Nusbaum, C., Lander, E., Abouelella, A., Allen, N.,
 Anderson, M., Arachchi, H. M., Barna, N., Bastien, V., Bloom, T.,
 Boguslavsky, L., Boukhgalter, B., Camarata, J., Chang, J., Choquel, Y.,
 Collymore, A., Cook, A., Cooke, P., Corum, B., DeRellano, K.,
 Diaz, J. S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
 Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S.,
 Graham, L., Grand-Pierre, N., Hages, B., Galagan, J., Gage, D.,
 Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
 Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R.,
 Lindblad-Toh, K., Liu, X., Lui, A., Mabbitt, R., Maclean, C.,
 MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
 Meldrum, J., Menes, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J.,
 Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
 O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
 Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
 Roman, J., Schauer, S., Schnepf, R., Seaman, S., Severy, P., Smith, C.,
 Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
 Talamas, J., Testafaye, S., Theodore, J., Topham, K., Travers, M.,
 Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X.,
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
 Direct Submission
 Submitted (12-FEB-2004) Whitehead Institute/MIT Center for Genome
 Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Feb 22, 2004 this sequence version replaced gi:11581779.
 All repeats were identified using RepeatMasker:
 Smit, A. F. A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L16015
Center clone name: 171_A_13

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 21715: contig of 21715 bp in length
* 21816 21815: gap of 100 bp
* 53733 53733: contig of 31918 bp in length
* 53734 53833: gap of 100 bp
* 53834 113500: contig of 59667 bp in length
* 113501 113600: gap of 100 bp
* 113601 196948: contig of 83348 bp in length
* 196949 197048: gap of 100 bp
* 197049 205662: contig of 8614 bp in length
* 205663 205762: gap of 100 bp
* 205763 235759: contig of 29997 bp in length.
Location/Qualifiers
1. .235759

FEATURES
source
1. .235759
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/map="14"
/clone="RP23-171A13"
/clone_lib="RPC1-23 Female Mouse BAC"

ORIGIN

Query Match 83.8%; Score 4804.8; DB 2; Length 235759;
Best local Similarity 92.0%; Pred. No. 0;
Matches 5316; Conservative 0; Mismatches 107; Indels 357; Gaps 9;

QY 1 GGATCTCGAAGTTCACAGAGGCTTCACCCACGAGTCCCTAGTTCATTTAGT 60
DB 211665 GGATCTCGAAGTTCACAGAGGCTTCACCCACGAGTCCCTAGTTCATTTAGT 211724
QY 61 TTCATGCTTTGTTCTCAATGCTGCTCCCGAGCTAATTTGATTTTAT 120
DB 211725 TTCATGCTTTGTTCTCAATGCTGCTCCCGAGCTAATTTGATTTTAT 211784
QY 121 TTCAAAAGGCTGTAATGAGAGTATCTTGTGCTACCCAGCTTAAAGGTGCCGTGA 180
DB 211785 TTCAAAAGGCTGTAATGAGAGTATCTTGTGCTACCCAGCTTAAAGGTGCCGTGA 211844
QY 181 AGCCCTGAGCTGAGGCTTTGCAACGCTTTAGTGTGAACAGATTAAGCAATTT 240
DB 211845 AGCCCTGAGCTGAGGCTTTGCAACGCTTTAGTGTGAACAGATTAAGCAATTT 211904
QY 241 TCCTTAAAGCCAAATCTGCTCTAGACTCTTCTCTGACCTCGTCCCTGCGCTCT 300
DB 211905 TCCTTAAAGCCAAATCTGCTCTAGACTCTTCTCTGACCTTGTGCTCCCTGCGCTCT 211964
QY 301 AGGCTGGGAGGTGGGCTTGAAGAGAGAGTGGGAGTGGCAAAAGCCGATCCCTAG 360
DB 211965 AGGCTGGGAGGTGGGCTTGAAGAGAGAGTGGGAGTGGCAAAAGCCGATCCCTAG 212024
QY 361 GGGCTGTGAAGTTCGAGGCTTCCCTGTAGACAGCTGGCTCATAGATCCCTCCGAGCC 420
DB 212025 GGGCTGTGAAGTTCGAGGCTTCCCTGTAGACAGCTGGCTCATAGATCCCTCCGAGCC 212084
QY 421 AAACATAGCAAGATGATACCTCTTTGTGACTTCCCAAGCCCAAGTACTGTCAAGTT 480
DB 212085 AAACATAGCAAGATGATACCTCTTTGTGACTTCCCAAGCCCAAGTACTGTCAAGTT 212144
QY 481 GAAACAGGATTTAGAAAGCTCTGAACCTCACTGAACCTGAAAGCTCATCCACCAAGCA 540

DB 212145 GAAACAGGATTTAGAAAGCTCTGAACCTCACTGAACCTGAAAGCTCATCCACCAAGCA 212204
QY 541 AGCACTAGATGCTCACTGCTAGTATGATTCATACGCTGATTAATATATGACAGAGTGGCCAC 600
DB 212205 AGCACTAGATGCTCACTGCTAGTATGATTCATACGCTGATTAATATATGACAGAGTGGCCAC 212264
QY 601 AGAAGTCTGGGGGTGAGAACTGACCACTGACTTTTCAGTGGGCAAGGTATGACCCCC 660
DB 212265 AGAAGTCTGGGGGTGAGAACTGACCACTGACTTTTCAGTGGGCAAGGTATGACCCCC 212324
QY 661 TCGAGATGTATATATGTCCTTATGATCCATCCAGGAGGTCTTAAAGGACAG 720
DB 212325 TCGAGATGTATATATGTCCTTATGATCCATCCAGGAGGTCTTAAAGGACAG 212384
QY 721 GGATGAGATGTATGATGATGAGCTTCCAAACAGCTATCCACAGTGTCCCTGCCCC 780
DB 212385 GGATGAGATGTATGATGATGAGCTTCCAAACAGCTATCCACAGTGTCCCTGCCCC 212444
QY 781 TTCACCTTAGCCAGAGAGCAAGTAACTTACCTTCTTCTCCCATCTCCAG 840
DB 212445 TTCACCTTAGCCAGAGAGCAAGTAACTTACCTTCTTCTCCCATCTCCAG 212504
QY 841 GACACACCCCTGCTGTGAGATTTTCAATTTCTTCTTCAAGTCCCTCTGTGACTTCCAT 900
DB 212505 GACACACCCCTGCTGTGAGATTTTCAATTTCTTCTTCAAGTCCCTCTGTGACTTCCAT 212564
QY 901 TTCGAAGGCTTTTGAAGCTGCTGAGCTGCTGGAAGATGATTTGGCTTAAAGTGTGACAA 960
DB 212565 TTCGAAGGCTTTTGAAGCTGCTGAGCTGCTGGAAGATGATTTGGCTTAAAGTGTGACAA 212624
QY 961 GCCATCTCAAGAAAGCAAGCAAGGGGAGCAGATTTTGAAGATTCAGAACTTAA 1020
DB 212625 GCCATCTCAAGAAAGCAAGCAAGGGGAGCAGATTTTGAAGATTCAGAACTTAA 212684
QY 1021 TCATGCGCGGCTGGGGGTGAGAAAAAGAGAGTATGCTCGCTCAGCTAAGCCAAAG 1080
DB 212685 TCATGCGCGGCTGGGGGTGAGAAAAAGAGAGTATGCTCGCTCAGCTAAGCCAAAG 212744
QY 1081 TAGTCCCGAGATATCTGCGCAGAGCTGGGCTGCTGGGGTATGTAAGATGTGGGTC 1140
DB 212745 TAGTCCCGAGATATCTGCGCAGAGCTGGGCTGCTGGGGTATGTAAGATGTGGGTC 212804
QY 1141 TGAAGCAATGGGATTTGAAGACATCTCTTTGATCTCCCTCAACCCCACTTACAGAC 1200
DB 212805 TGAAGCAATGGGATTTGAAGACATCTCTTTGATCTCCCTCAACCCCACTTACAGAC 212864
QY 1201 ACACTGCTGTGGCCAGACTCTCTGTTCACAGCCCTGCTGTGACCACTGAGCTAG 1260
DB 212865 ACACTGCTGTGGCCAGACTCTCTGTTCACAGCCCTGCTGTGACCACTGAGCTAG 212924
QY 1261 GCAACGAGCATGGGCTCTGTGCTGAGATGAAAGTTGTTACCAATAGCAAAAAACAG 1320
DB 212925 GCAACGAGCATGGGCTCTGTGCTGAGATGAAAGTTGTTACCAATAGCAAAAAACAG 212984
QY 1321 CAGGGAGGAGAAACAGAAAGCAATTAAGAAAGAAAGAAAGGCACTCAATCAGA 1380
DB 212985 CAGGGAGGAGAAACAGAAAGCAATTAAGAAAGAAAGAAAGGCACTCAATCAGA 213044
QY 1381 TGCATGTCAAGAGATGGAAGCCCAACACACACTTGAAGAGGAAACAGAAAGGAG 1440
DB 213045 TGCATGTCAAGAGATGGAAGCCCAACACACACTTGAAGAGGAAACAGAAAGGAG 213104
QY 1441 AGATTCTGGGATTAAGAGGCTCAGAAAGAGAGCCAGCCCAAGTCTCTCTTT 1500
DB 213105 AGATTCTGGGATTAAGAGGCTCAGAAAGAGAGCCAGCCCAAGTCTCTCTTT 213164
QY 1501 ATACCTCATCCGCTCTCCCAATTAAGCCACTTCTTCTTATGATCAGACTGAGCTGC 1560
DB 213165 ATACCTCATCCGCTCTCCCAATTAAGCCACTTCTTCTTATGATCAGACTGAGCTGC 213224
QY 1561 AGGGAAGACCCCTAGGAGATCACTGTATGAAAGAGTGTGGGAAAGTCCAGG 1620

Db	213225	AGCGAAGAGACCCGTAAGGAGATCAACTGGATGAAGAGATGTGTGGAGAAAGTCCAG	213284
OY	1621	GCACCTTAAGAGCCAGAGCCTTAAGAAAGCAAGATTAAGGTGTCTTCAAGAGTGCCAG	1680
Db	213285	G-AACTTAAGAGCCAGAGCCTTAAGAGCAAGATTAAGGTGTCTTCAAGAGTGCCAG	213433
OY	1681	CTGTGCACACAGAGGCTCGAGGACTGTGTGTAGAGCTTCAAGATTAAGATGTGATGCTCA	1740
Db	213344	CTGTGCACACAGAGGCTCGAGGACTGTGTGTAGAGCTTCAAGATTAAGATGTGATGCTCA	213403
OY	1741	ATGGCGGGGGG-GGGGATTTCTGGGGGGGGAGAGAAAGTGTGAAGAGAGCCTGGAAC	1799
Db	213404	ATGGCGGGAGGTGGGGGAGTCTTGGGGGGAGAGAAAGTGTGAAGAGAGCCTGGAAC	213463
OY	1800	GAGAACTTGAAGCCCTGTGAAAAGATACCTAAGGGAAAGAAACCAAGCTACTTTAGAT	1859
Db	213464	GAGAACTTGAAGCCCTGTGAAAAGATACCTAAGGGAAAGAAACCAAGCTACTTTAGAT	213523
OY	1860	GTAATTCATGAAAGACAGGGAGAGAGAAAGCTGTGAGAGAGT--AGAAAGACCCCGGGG	1916
Db	213524	GTAATTCATGAAAGACAGGGAGAGAGAAAGCTGTGAGAGAGTGTGAAGAAAGAACCCCGGGG	213583
OY	1917	AAGACATGGAGACAAAGAACCAAGCTTGAAGCGCTCCGTGAATTAAGCTGTGAAGG	1976
Db	213584	AAGACATGGAGACAAAGAACCAAGCTTGAAGCGCTCCGTGAATTAAGCTGTGAAGG	213643
OY	1977	AGAGCCCTGTGTGTGAGCACCAAGACAGCAGAGGCTTAGGTAAATGTCTGAGACAGGAA	2036
Db	213644	A-----GAGCACCAAGAACCAAGAGCTTAGGTAAATGTCTGAGACAGGAAAC	213691
OY	2037	GAAAGTAGACACAGGAAACAGACAGAGACGGGGAGCCAGGTAACAAGAAATGTCTTTC	2096
Db	213692	GAAAGTAGACACAGGAAACAGACAGAGACGGGGAGCCAGGTAACAAGAAATGTCTTTC	213751
OY	2097	TCACCTGTGGCCAGAGCGTCCATCTGTGTCCATCTCTTAAGATGTTCATCAGATCTGA	2156
Db	213752	TCACCTGTGGCCAGAGCGTCCATCTGTGTCCATCTCTTAAGATGTTCATCAGATCTGA	213811
OY	2157	GGGCTGTGCTTGGAGAGCAGCTGGAAGAGTATGTGTGAGCCAGGGAGACAAAGGGGGCT	2216
Db	213812	GGGCTGTGCTTGGAGAGCAGCTGGAAGAGTATGTGTGAGCCAGGGAGACAAAGGGGGCT	213871
OY	2217	AGGAAGGAAGAAAGAGGGCAAAACAGGCCACAAAGGGCAGAGCCCAAACTGAGTTA	2276
Db	213872	AGGAAGGAAGAAAGAGGGCAAAACAGGCCACAAAGGGCAGAGCCCAAACTGAGTTA	213931
OY	2277	ACTCTTCTCTTGTGATTTTCATATGAGAGCAGTGTGAAACTCTGTGACCACTATCCCC	2336
Db	213932	ACTCTTCTCTTGTGATTTTCATATGAGAGCAGTGTGAAACTCTGTGACCACTATCCCC	213991
OY	2337	ATGAGCCCCCACTAACCATACCAAGTTTGGCCGATGTGCACTTCTAGGTTCCCTGAGAC	2396
Db	213992	ATGAGCCCCCACTAACCATACCAAGTTTGGCCGATGTGTGCACTTCTAGGTTCCCTGAGAC	214051
OY	2397	AGAGCTGTGCTTTGTCTCTTGTGACCTGACCCAAAGCTGACCCCAATGTTCTCAGTACCTT	2456
Db	214052	AGAGCTGTGCTTTGTCTCTTGTGACCTGACCCAAAGCTGACCCCAATGTTCTCAGTACCTT	214111
OY	2457	GAATGCGCTCAAGAGCTTGAGAAACAGGCAAGTGAATATTAGGCAATGGGCTTAACCTGG	2516
Db	214112	TCATGCGCTCAAGAGCTTGAGAAACAGGCAAGTGAATATTAGGCAATGGGCTTAACCTGG	214171
OY	2517	AGCTTGACACAGAGGCGTCAAGTGAACCTCAAGGGACACAGCTGACAGACAGGTGGGCTTT	2576
Db	214172	AGCTTGACACAGAGGCGTCAAGTGAACCTCAAGGGACACAGCTGACAGACAGGTGGGCTTT	214231
OY	2577	ATCCCAAAAGACAACCATTTGGCATAAGTGTGCAAAATGGGAATGCAAAGTTTGATCA	2636
Db	214232	ATCCCAAAAGACAACCATTTGGCATAAGTGTGCTCAAAACGGGAATGCAAAGTTTGATCA	214291
OY	2637	GGTCCCTTCAAGATATCTGCATGACAAAGACCTTAAGACCCCTTGAGAGAGGGGTATGTCTCT	2696
Db	214292	GGTCCCTTGAATAATCTGCATGACAAAGACCTTAAGACCCCTTGAGAGAGGGGTATGTCTCT	214351

OY	2697	GCCTCCACCCACCATTAAGGGGAGTGAATCTATCTTGGGGAGCTGGGACACTTGGGGAGACA	2756
Db	214352	GCCCCCACCCACCAATMAAGGGAGTGMAATCTCTGAGGGGAGCTGGGACACTTGGGGAGACA	214411
OY	2757	CCACATTAAGAGAGTGTGAGAGCCAGAGAAAACTAGCCGCCCTGTGTCTGGCCCACTCC	2816
Db	214412	CCATATTACTAGAGTGTGAGAGCCAGAGAAAACTAGCCGCCCTGTGTCTGGCCCACTCC	214471
OY	2817	ACACTTAAGACTATATTTGAAGGTGACAGTAGATAGTAGGGTGGAGCTGGTACAGGGAGA	2876
Db	214472	ACACTTAAGACTATATTTGAAGGTGACAGTAGATAGTAGGGTGGAGCTGGTACAGGGAGA	214531
OY	2877	GTGTTTCCGGGGTGGAGGGGTGATAGGGGAAAGCCAGAGAGGGGAGTCTGGCTTTGTCTCC	2936
Db	214532	GTGTTTCCGGGGTGGAGGGGTGATAGGGGAAAGCCAGAGAGGGGAGTCTGGCTTTGTCTCC	214591
OY	2937	TGAACACATGTCTACTTATGTTATTAACAGCATGACCTTGCTAAGAAGCCCAACATCTACGA	2996
Db	214592	TGAACACATGTCTACTTATGTTATTAACAGCATGACCTTGCTAAGAAGCCCAACATCTACGA	214651
OY	2997	CCTCTGAAAAAGACAGACAGCCCTTGGAGGACAGGGGGTTGTCTCTGAGCCTTGGGTCTTGAT	3056
Db	214652	CCTCTGAAAAAGACAGACAGCCCTTGGAGGACAGGGGGTTGTCTCTGAGCCTTGGGTCTTGAT	214711
OY	3057	GGTGCCACAAAGAGAGGGATGAGTGTGATTAAGAGCCCCAGAGAGGCTTAAGAAAGGCA	3116
Db	214712	GGTGCCACAAAGAGAGGGATGAGTGTGATTAAGAGCCCCAGAGAGGCTTAAGAAAGGCA	214771
OY	3117	CTTGGGAAAGGGGTGAGTGTGACAGAGCCCTTATCCATGGAATGTGAGAGCTGGGGCCAAT	3176
Db	214772	CTTGGGAAAGGGGTGAGTGTGACAGAGCCCTTATCCATGGAATGTGAGAGCTGGGGCCAAT	214831
OY	3177	GGTGTAAATCTCTGGGCGCTGCGACGGCATTTCAAGAGACACCTGCATCCCTGGACGCTG	3236
Db	214832	GGTGTAAATCTCTGGGCGCTGCGACGGCATTTCAAGAGACACCTGCATCCCTGGACGCTG	214891
OY	3237	GGGAGGCGGAAAGGAGCAACCCCACTTATACCCCTTCTCCCTCAGGCCACAGATTAAC	3296
Db	214892	GGGAGGCGGAAAGGAGCAACCCCACTTATACCCCTTCTCCCTCAGGCCACAGATTAAC	214951
OY	3297	ACCTCTGGCCTTTCCCCCTTCCCACTCCCATAGAGAGTGAAGGGTTGGAGAGGAGGGTA	3356
Db	214952	ACCTCTGGCCTTTCCCCCTTCCCACTCCCATAGAGAGTGAAGGGTTGGAGAGGAGGGTA	215011
OY	3357	AAAACTCATGTCCAAACATCATGTGTGCAACAATTAATGATCAGTATGTGTAGAGGCAA	3416
Db	215012	AAAACTCATGTCCAAACATCATGTGTGCAACAATTAATGATCAGTATGTGTAGAGGCAA	215071
OY	3417	GAAGGAATCTGACAGGCTTAACTGGGTTAATGTGTAAAGTGTGTGTGATGTGTGTG	3476
Db	215072	GAAGGAATCTGACAGGCTTAACTGGGTTAATGTGTAAAGTGTGTGTGATGTGTGTG	215131
OY	3477	TCTGACGTAAAGACGGGCAATGGGTGTGACGTGTGATCTGTGGGTGAGGTTTACAGAC	3536
Db	215132	TCTGACGTAAAGACGGGCAATGGGTGTGACGTGTGATCTGTGGGTGAGGTTTACAGAC	215191
OY	3537	TGCAGGTTTGTGTGTAAATTTGCCCAAGGCAAAAGGGGTGAATCCCTTCCATGGTTTAAAG	3596
Db	215192	TGCAGGTTTGTGTGTAAATTTGCCCAAGGCAAAAGGGGTGAATCCCTTCCATGGTTTAAAG	215251
OY	3597	AGATTGGATGATGGCCTGTGATTTCAAGACCATTGGAAAAATAGATGGACACTTATATGT	3656
Db	215252	AGATTGGATGATGGCCTGTGATTTCAAGACCATTGGAAAAATAGATGGACACTTATATGT	215311
OY	3657	GTCTCTAAGCTAAAGGTAGCAAGGCTCTTTGAGAGACACCTGTCTAAGATGTGGGCAACAG	3716
Db	215312	GTCTCTAAGCTAAAGGTAGCAAGGCTCTTTGAGAGACACCTGTCTAAGATGTGGGCAACAG	215371
OY	3717	AGACTTACAGACGATCTCTGTACAGAGTAGAGAGAGAGAGAGGGG-GTGTAGAAATCTCT	3775
Db	215372	AGACTTACAGACGATCTCTGTACAGAGTAGAGAGAGAGAGAGGGGCGCTAGAAATCTCT	215431

QY	3776	TACATCAAAAGGAAAC	TGAGTCGCGAC	CTGCAAAAGTGAAG	CTCTCCCTAGAC	TCAAT	3835
Db	215432	TACTATCAAAAGGAAAC	TGAGTCGCGAC	CTGCAAAAGTGAAG	CTCTCCCTAGAC	TCAAT	215491
QY	3836	GACTTGTCTCTGGGAG	CGACACACTGTGAA	CTTCAGGTCTGAGA	GTAGAGGCTCC		
Db	215492	GACTTGTCTCTGGGAG	CGACACACTGTGAA	CTTCAGGTCTGAGA	GTAGAGGCTCC		215551
QY	3896	CCTCAGCCTGAAAGCT	ATGCAATAGCCAGGG	TTGAAAGGGGGAAG	GAAGACCTGGGATG	3955	
Db	215552	CCTCAGCCTGAAAGCT	ATGCAATAGCCAGGG	TTGAAAGGGGGAAG	GAAGACCTGGGATG	215611	
QY	3956	GGAGCTTGATGTGAG	GCAGGGGACAGAT	TTAAGCTCGAAGAA	GAAGGTGACCCCTTA	4015	
Db	215612	GGAGCTTGATGTGAG	GCAGGGGACAGAT	TTAAGCTCGAAGAA	GAAGGTGACCCCTTA	215671	
QY	4016	CCCAATGTTCAACTC	ACCCCTTCAGATTAA	AAAAATACTAGAGT	TAAGGGCTGGGTAGG	4075	
Db	215672	CCCAATGTTCAACTC	ACCCCTTCAGATTAA	AAAAATACTAGAGT	TAAGGGCTGGGTAGG	215731	
QY	4076	AGGTGCTGTGAGAC	GCCTCTGTCTCTCT	GTGATGCTCCCTGA	GGCCCTTTGGGAGAG	4135	
Db	215732	AGGTGCTGTGAGAC	GCCTCTGTCTCTCT	GTGATGCTCCCTGA	GGCCCTTTGGGAGAG	215791	
QY	4136	AATGAGCCCAAGAC	TAATAAAAAAGCCAT	GGAGCCAGAGGGGGA	AGGGCAACAGACCTTT	4195	
Db	215792	AATGAGCCCAAGAC	TAATAAAAAAGCCAT	GGAGCCAGAGGGGGA	AGGGCAACAGACCTTT	215851	
QY	4196	CATGGGCAAACTTTG	GGGGCCCGTAGTGA	TCGATTGAACAAGAC	CTCCCAATCGATACCT	4255	
Db	215852	CATGGGCAAACTTTG	GGGGCCCGTAGTGA	TCGATTGAACAAGAC	CTCCCAATCGATACCT	215911	
QY	4256	TCTTCTTCTTAAC	GCAGCAGAGGGA	ACTCGAGTTTACCA	CTCCCTATCAGTATAG	GAAGA	4315
Db	215912	TCTTCTTCTTAA	CGACAGAGGGA	ACTCGAGTTTACCA	CTCCCTATCAGTATAG	GAAGA	215946
QY	4316	AGTGAAGTCGAGTT	TACCACTCCCTAT	CATAGTATAGAAAA	AGTGAAGTCAAGTTTAC	4375	
Db	215947	AGTGAAGTCGAGTT	TACCACTCCCTAT	CATAGTATAGAAAA	AGTGAAGTCAAGTTTAC	215947	
QY	4376	CACCTCCCTATCAG	TATAGAAAAAGTGA	AGTCGAGTTTACCA	CTCCCTATCAGTATAG	4435	
Db	215947	CACCTCCCTATCAG	TATAGAAAAAGTGA	AGTCGAGTTTACCA	CTCCCTATCAGTATAG	215946	
QY	4436	GAGAAAGTGAAAG	TGCGAGTTTACCA	CTCCCTATCAGTATAG	AAAAAGTGAAGTCGA	4495	
Db	215947	GAGAAAGTGAAAG	TGCGAGTTTACCA	CTCCCTATCAGTATAG	AAAAAGTGAAGTCGA	215946	
QY	4496	GTTTACCACTCCCT	ATCAGTATAGAAAA	AGTGAAGTCGAGTTT	ACCACTCCCTATCA	4555	
Db	215947	GTTTACCACTCCCT	ATCAGTATAGAAAA	AGTGAAGTCGAGTTT	ACCACTCCCTATCA	215946	
QY	4556	GTGATAGAGAAAA	GTGAAGTCGAGCT	CGGTACACAGAGAG	CTCCAAATTTAGCAGC	4615	
Db	215947	GTGATAGAGAAAA	GTGAAGTCGAGCT	CGGTACACAGAGAG	CTCCAAATTTAGCAGC	215979	
QY	4616	AGGCATATGAGAT	TGGATTTAAAGGG	CTTGAGCACTGAGA	GTGT---CAGAGATTTTC	4671	
Db	215980	AGGCATATGAGAT	TGGATTTAAAGGG	CTTGAGCACTGAGA	GTGT---CAGAGATTTTC	216039	
QY	4672	TCCAACCCAGGTTA	AGGAGATTTGGGTG	GGGGGCTTTCACCA	CAACCAAGACCTCTCCC	4731	
Db	216040	TCCAACCCAGGTTA	AGGAGATTTGGGTG	GGGGGCTTTCACCA	CAACCAAGACCTCTCCC	216099	
QY	4732	CACCTAGAGAGAA	ACTGCCCTTCTCTG	GAAGTGGGTTCAAG	CCGGTCCAGAGCTTGACAG	4791	
Db	216100	CACCTAGAGAGAA	ACTGCCCTTCTCTG	GAAGTGGGTTCAAG	CCGGTCCAGAGCTTGACAG	216155	
QY	4792	GGTGCCCTTCCACA	CCACTGAGAGTTTCA	GTGCGAGAGGATTTT	CCACAAGAAACA	CTG	4851
Db	216160	GGTGCCCTTCCACA	CCACTGAGAGTTTCA	GTGCGAGAGGATTTT	CCACAAGAAACA	CTG	216216
QY	4852	GATGCCCTTCCCTTA	AGCTGTCTTCCAT	CTTCTCTCGGGAGAT	CTCTCCCGCTT	4911	

Db	216217	GATGCCCTTCCCTTACGCTGTCTCTCATCTTCTCTCTGGGAGATGCTCCCTCCCGTCT	216276
Qy	4912	TGGTTTATCTTGGCTCTTGTCTTTCAGCAAGATTGGCCCTGTGTGTCCATCTT	4971
Db	216277	TGGTTTATCTTGGCTCTTGTCTTTCAGCAAGATTGGCCCTGTGTGTCCATCTT	216336
Qy	4972	CTTCACTGTCTCCGAGCCCTTGTGCTTGGCTCTTGGGTGTGCTCTTCCATCCCACTTCT	5031
Db	216337	CTTCACTGTCTCCGAGCCCTTGTGCTTGGCTCTTGGGTGTGCTCTTCCATCCCACTTCT	216396
Qy	5032	CACCTTCACTTCTTCCCTCTCTCATTTGTATTCATCTCTCTCTCTT-----	5083
Db	216397	CACCTTCACTTCTTCCCTCTCTCATTTGTATTCATCTCTCTCTCTT-----	216456
Qy	5084	-----CCTTCTCTCTTCTCTTCTCTTCTT	5111
Db	216457	CCTTCTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTTCTCTT	216516
Qy	5112	CCCTTCTCCCT	5171
Db	216517	CCCTTCT	216576
Qy	5172	TGCTGAGATCACACCTGGGGTCCACCCCTATATTAACAATCTTCCAGTAGCCACAG	5231
Db	216577	TGCTGAGATCACACCTGGGGTCCACCCCTATATTAACAATCTTCCAGTAGCCACAG	216636
Qy	5232	CTTCACTGTCTGTGGGAGCTCTTCACTTCTCTCACCCCCTGGCTTGTCTCTCATCC	5291
Db	216637	CTTCACTGTCTGTGGGAGCTCTTCACTTCTCTCACCCCCTGGCTTGTCTCTCATCC	216696
Qy	5292	TGTGTCAAGATCTTCAATTTGGTGTCTCCAGCCCTCTGTACTCTCTTCCGTGCTTCTTC	5351
Db	216697	TGTGTCAAGATCTTCAATTTGGTGTCTCCAGCCCTCTGTACTCTCTTCCGTGCTTCTTC	216756
Qy	5352	TCTGTGTCAAGTGGCCACTGTGTGTGCTGTTCAGCTGTGTGTCCACATTTCTCAGA	5411
Db	216757	TCTGTGTCAAGTGGCCACTGTGTGTGCTGTTCAGCTGTGTGTCCACATTTCTCAGA	216816
Qy	5412	TTCTCTGAAAAGTTAACCAAGGTAGAAATGTTTCCCTGTAGACAGACATCAAGATTCTC	5471
Db	216817	TTCTCTGAAAAGTTAACCAAGGTAGAAATGTTTCCCTGTAGACAGACATCAAGATTCTC	216876
Qy	5472	CCGGAAGTCAAGCTTCCAGCCCTCTCTTTCTGTGCCAGCTGCCGGCACTCTTACAAA	5531
Db	216877	CCGGAAGTCAAGCTTCCAGCCCTCTCTTTCTGTGCCAGCTGCCGGCACTCTTACAAA	216936
Qy	5532	CCTGAGGACCCCTTACCCCACTATAGCCCTGTACAGAGAGAGGACCTTACATGAGT	5591
Db	216937	CCTGAGGACCCCTTACCCCACTATAGCCCTGTGTACAGAGAGGACCTTACATGAGT	216996
Qy	5592	CCTGTGTGAGAGCCATAGGCTACGAGTGTAAAAAGAGGACGGGAAGTGTGTAGAAA	5651
Db	216997	CCTGTGTGAGAGCCATAGGCTACGAGTGTAAAAAGAGGACGGGAAGTGTGTAGAAA	217056
Qy	5652	GTCAGGACTTTCATATGAAGCTTACGCCACACCAAGAAATGACAGACAGATCCCTTATC	5711
Db	217057	GTCAGGACTTTCATATGAAGCTTACGCCACACCAAGAAATGACAGACAGATCCCTTATC	217116
Qy	5712	TCCCCCATTAAGATTGAGT 5731	
Db	217117	TCCCCCATTAAGATTGAGT 217136	
RESULT 4			
LOCUS	AC130940	185702 bp	DNA linear HTG 15-NOV-2002
DEFINITION	Rattus norvegicus c1one CH230-249H16, WORKING DRAFT SEQUENCE.		
ACCESSION	AC130940		
VERSION	AC130940.3	GI:25007324	
KEYWORDS	HTG; HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		

Db 97260 TCCCTTCAAGCCAAATCTCTCTCTAGACTCTTCTTCACTGACCTCGCTCCGCGGCTC 97319
Qy 300 TAGGCTGGGAGGTGGGCTTTGGAAGAAAGGTGGGAGTGGCAAAAGCCGCTCTA 359
Db 97320 TAGGCTGGGAGGTGGGAGCTTTG--AAAGAAAGGAGGAGAGTGGCAAAAGCCGCTCTA 97376
Qy 360 GGGCCCTGTGAAGTTGAGAGCTTCCCTGTGACAGACTGGCTCATAGATCTCTCCAGC 419
Db 97377 GGACTGTGAAAGTTGTAGAGCTTCCAGTACAGTGTGGCTCATAGATCTCTCCAGC 97436
Qy 420 CAAACATAGCAAGAAAGTATACCTCTTTGTGATTTCCCAAGGCCAGTACCTGTCAAGT 479
Db 97437 CAAACACAGCAAGAAAGTATGCTTCTTTGTGATTTCCCAAGGCCAGTACCTGTCCGCT 97496
Qy 480 TGAACACAGATTTAGAGAGGCTCTGAACCTCACTGAACCTGTGAAGCTATCCACCAAGC 539
Db 97497 TGAACACAGATTTAGAGAGGCTCTGAACCTCACTGAACCTGTGAAGCTATCCACCAAGC 97552
Qy 540 AAGCAGCTAGAGTGCCTGCTAGTATGATCTTACGCTGATTAATATGACAGAGCTG--GCG 597
Db 97553 ACACAGCTTACATGCGGGTCTTAATCAGATCTTCACTGATTAATACGACAGCTGACCAAC 97612
Qy 598 CACAGAGTCTGTGGGCTGTAGAACTGACCAAGTACCTTTTCACTGCGCAAAAGGTATGACC 657
Db 97613 GGGAGAAAGGCTGAGTGTAGGAACGTGACAGTACCTTTTCACTGCGCAAAAGGTATGACC 97672
Qy 658 CCCCTCAGCAGAT---GTAGTAATGCTCCCTTAGATCCCATCCAGGAGGCTCTTAAGAG 714
Db 97673 CCCCTGCAAGTATGAGCAGATGGGTCCCTTAGAACCAATCCAGGAGGCTCTGTAGAG 97732
Qy 715 GACATGAGATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774
Db 97733 GACACAGATGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 97792
Qy 775 TGCCCCCTCCCTTAGCAGAGAGAGACAGTAACTTTTACCTATCTTTTCTCTCCCATCC 834
Db 97793 TGCTCCCTTCACTTAGCAGATGAGATGAGCACTTTACCTAATCTTTTCTCTCCCATCT 97852
Qy 835 TCCAGAGACACACCCCTGTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 894
Db 97853 TCCAGAGACCAACCCCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 97911
Qy 895 TTCCATTTGCAAGGCTTTTGACTCTGCAAGCTGCTGGAAGATAGAGTTTGGCCCTAGGTG 954
Db 97912 TTCCATTTGCAAGGCTTTTGACTCTGCAAGCTGCTGGAAGATAGAGTTTGGCCCTAGGTG 97971
Qy 955 TGGCAAGCCCTTCTGAAGAAAGACAGACAGAGGAGACAGATTTTGGAAAGATCAAGCA 1014
Db 97972 TGGCAAGCCATCTTGAAGAAAGACAGACAGAGGAGACAGATTTTGGTATAGG--GGGAC 98030
Qy 1015 ACTTAATCACTGCGGCGCTGGGGTAGAAAAAAGAGTGAAGTGCCTCCAGCTAAG 1074
Db 98031 ACTTAATCACTGCGGCGCTGGGGTAGAAAAAAGAGTGAAGTGCCTCCAGCTAAG 98082
Qy 1075 CCAAGTATGCTCCGAGATCTGCAACAGCTGGGCTGCTCGGGTAGCTTTAGAGATG 1134
Db 98083 CAAAGCAGATCTTGGGGTATTT---CAAGCTGGGCTGCT--GAGGTATGATTTAGAGATG 98137
Qy 1135 TGGGTCTGAAGAAAGATGGGATTTGGAAAGATCTCTTTGAGTCTCCCTCAACCCCACT 1194
Db 98138 TGGGTCTGAAGAAAGATGGGATTTGGAAAGATCTCTTTGAGTCTCCCTGCTTGGGCT 98197
Qy 1195 ACAGACACACTGCTGTGTGCGCAGACTCTGTTCAACAGCCCTCTGTGTTTGAACACTG 1254
Db 98198 AAATCTCCCTCAACCCCACTATGAC-----ACGCTGTGATGAGCCAGACTG 98247
Qy 1255 AGTAAAGCAACAGAGATGGGCTCTGTGCTGAGATGAAGATTTGTTTCAATATGACAA 1314
Db 98248 AGTAAAGCACTGAGATGGGCTCTGTG--TGAGGACGAG-----GGTTGCAATATGACAA 98302
Qy 1315 AAACAGCAGGAGGAGGAGAAACAGAGAACGAATTAAGAAAGAAAGAAAGGCGCACTCA 1374

Db 98303 AAACAGCAGTGAAGGAGAAACAGAAATTAAGAAAGAAAGACGAAACGCGC---CA 98359
Qy 1375 ATCAGATGCAAGTCAGAAAGATGGAAGGCCAAACACAGCTTGAAGCAGAAACAGAA 1434
Db 98360 ATCAGATGCAAGTCAGAAAGATGGAAGAACCAATGCTTACTTGAACAAAGAAATTAAGAG 98419
Qy 1435 AGGAGAGAG----- 1443
Db 98420 AGGAGAGATGAGCAGATCATTAAGGTGAGCTTCTCATTAATTAAGATGCTCTGTGT 98479
Qy 1444 -----TTGCGGCATTAAGGAGGCCACAGAAAGAGGCCAGGCC 1484
Db 98480 CCGACTTCAATCCCTGGGTAAGTCTTGTGCACTCAGAGATTAAGCAGACAGAAACT 98539
Qy 1485 CCCAAGTCTCTCTTTATATCCCTCATCC----- 1513
Db 98540 GTATATGGCATTTTCAATTCATTCGTAAGCTTGAACACTTGACACTTGAATTTTATGTT 98599
Qy 1514 ----- 1513
Db 98600 TACATTTGTATCTCTATACAGATATGCTTCAATGAGCAGTGTGATGCTCTGAAAT 98659
Qy 1514 ----- 1513
Db 98660 ACAAGTCTTTTATTAAGAAATCAAGTGCATTAATCAAGTCTGATAGGAGGCCA 98719
Qy 1514 -----GTCTCCCAATTAAGCCCACTCTTCTCTAGATCAGAC 1551
Db 98720 CAGGAAGAGCCAGATCCCGAGCTCTCTTTATGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 98779
Qy 1552 CTGAGCTGCAAGCAAGAGACCCGTAGAGAGATCACTGATGAAGAGATGTGTGAGAG 1611
Db 98780 CTCTTC--CTGGAACACACTGAGCAGCAGTGTGAGAAATCCGTAGACTTAAGAGATCA 98838
Qy 1612 AAGTCCAGGCAACCTTAAGAGCCAGAGCTTAAGAGAGAGATTAAGAGTCTTCAAG 1671
Db 98839 CAGTGAAGAGATGTGTGAGCAGAGCTTAAGAGAGAGATTAAGAGAGCAGCAAG 98898
Qy 1672 GTGGCAGGCTGTG--CACACAGAGGTCAGAGACTGTGTGAGAGCTTCAAGATTAAGCA 1729
Db 98899 GCACTTATGAGCCGTGACACACAGAGGTCAGAGACTTATGAGGTCCTCAAGATTAAGCA 98958
Qy 1730 TGATGCTCAGATGAGGCGGGGGGGGGGATTTCTGGGGGGGGAGAGAGAGTGAAGAGCA 1789
Db 98959 TGATGCTCAGATGAGGCGGGGGGGGGGAGAGC----- 98987
Qy 1790 GCCTGGAACAGAGATCTGGAAGCGCTGGAAGCATACCAATAAGGAGAAACCCAGGCT 1849
Db 98988 -----GGAGCTGAGACTGTGAGAGCGCTGGAAGCATACCTTTAAAGGAGAAACCCAGGCT 99043
Qy 1850 ACCTTATGATGTAATCATGAAAGACAGAGAGAGAGAGAGCTTGAAGATGAAGAGACC 1909
Db 99044 ACCTTATGATGTAAGTCAATGAAGTCAAGGAGAGAGAGAGAGCTTGAAGATGAAGAG----- 99099
Qy 1910 CCGAGGCAAGACATGAGAGCAGAGCAGGCTTGAAGCGCTCCGTGAATCAAGCTGTC 1969
Db 99100 -----AATTAAGCAAGAGCAGAGCTTGAAGCGCTCTGTGAAT----- 99139
Qy 1970 TGAAGCAGAGCCCTGTATGAGCAGAGAAACAGAGAGCTTGAAGCTTGAAGCA 2029
Db 99140 -----TCCATATGAGCGCGGAGAACAGAGAGCTTGAAGCTTGAAGCA 99187
Qy 2030 GGGAGAGAGATGAGCAGAGAAACAGAGACAGAGAGCGGGAGCCAGTAAACAAAGAGATG 2089
Db 99188 GGGAGAGAGATGAGGAGCAGAGAAACAGAGACAGAGAGCGGGAGATCAAGGTAAACAAAGAGATG 99247
Qy 2090 GTCTTCTCACTGTGTGCGCAGAGCTTCATCT--GTGTCCATATCTTGAAGATTTCAATC 2148
Db 99248 GTCTCTCTCACTGTGTGCGCAGAGCTTCATCTGTGGGTTCATCTTGAAGATTTCAATC 99307
Qy 2149 AGACTGAGAGGCTGTGTGAGAGCAGCTGAAGAGATGTGAGAGCCAGGAGAGAGCA 2208
Db 99308 AGACTGAGAGGCTGTGTGAGAGCGGCTGAGAGAGATGTGAGAG--CGGGGAGAGCTA 99366

Db	101421	-----	101420
Oy	4412	GTTTACCACTCCCTTACAGTATAGAGAAAAGTAAAGTCAGTTTACCACTCCATCA	4471
Db	101421	-----	101420
Oy	4472	GTGATAGAGAAAAGTAAAGTCAGATTTACCACTCCCTATCAGTATAGAGAAAAGTAA	4531
Db	101421	-----	101420
Oy	4532	AGTCAGTTTACCACTCCCTATCAGTATAGAGAAAAGTAAAGTCAGCTCGTACAG	4591
Db	101421	-----	101428
Oy	4532	CAGAGACTCCAAATTTAGGACAGACAGGCATATGGATGTAAAGGGGCTGAGACAC	4651
Db	101429	CAGATGACTCCAAATTTAGGACAGACAGGCATGTGAATGTACTAAAGGGGCTGAGACCC	101488
Oy	4652	TGAGAGCTGTCA-----GAGATTCTTCCAAACCAAGGTAGAGGAGATTTCGGGTGGGGGCT	4707
Db	101489	TGAGAGCTGTCAAGACCGAGATTTCTTCATCCAGGTAAAGAGATTTACGTGGGGGCT	101548
Oy	4708	CTTCCACCAACACAGACTCTCCCACTAGAGAGAACTGCCTTTCCTGGAATGGGGGT	4767
Db	101549	CTCCAAACGACACAGACCTGTCCCACTAGAGAGAAAGTGTCTTCCGGAAGTGGGGG	101608
Oy	4768	TCAGGCGCGGTCAAGATCTGAACAGGGTGGCTTCCACAGCTCGGGAAGTTCTCACTGGC	4827
Db	101609	T-----CTTCCACCGGGCTGGGAAGTTCCTGGTGGC	101640
Oy	4828	AGAGAGTTTCCACAAGAAACACTGATGAGCCCTCCCTTACGCTGTCTTCATCTTCC	4887
Db	101641	AGGATGTTTC-----TACTGATGACCTTCCCTTCCACTGCTCTCTCCCTC----	101687
Oy	4888	TCCTGGGAGATGCTCCTCCCGTCTTGGTTATCTTGGACTCTTGTCTTCAAGCAAGATTG	4947
Db	101688	-----CTTGTCTTATTAATCTTGGCTTCTTATGTTTCAAGAAAGATTGG	101731
Oy	4948	CCCTGTGCTGTCCACTCCCATCTTTCCTTACTGTCTCCGAGCTTGGCTTGGCTTTGGG	5007
Db	101732	CCCGGTGCTGTACTCCACTGTCTCTTACTCTCTC-----TGCCTGGCTTTCTGTGG	101784
Oy	5008	TGTCCTTCTTTCACACCAATTTCTCACTTCACTTTCCTCCCTTCTCATTTGTATTCAAT	5067
Db	101785	TGTTCTCTTTTTCAGAGTGTTTCTCACTCACTGCTCCCGCCCGCTTCATTTTTA	101843
Oy	5068	CCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	5127
Db	101844	-----	101843
Oy	5128	CCTTCCCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT	5187
Db	101844	-----	101884
Oy	5188	TGGGGTTC--CACCTTATGTAAACAATCTTCCAGTAGACCAAGCTTCAGTGCCTG	5245
Db	101885	CAGGGCTTATACAGTCAAGTGAAGATCTCCCACTGAGCCAAAGCTTTAATCT-CTG	101943
Oy	5246	GGTGTCTCTTACCTTCTTCAACCCCGTGGCTTGTCTTTCATCTCTGTGACGAATCTCT	5305
Db	101944	GGTGTGTCTTACCGACACTCACTCCCTGTCTTCTTCTTCCGCTCTAAGACGAATCTCT	102003
Oy	5306	AGATTGTCTCCAGAGCTCTGTACTACTCTTCTTCTGCTGTCTCTCTTGTTCACGTG	5365
Db	102004	-GGTCCGTCTTACAGTCTTGTCTACTACTCTTCTTCTGCTGTCTTCTTCTCCGTCCAGCTG	102062
Oy	5366	CGCACACTGTGTGCTCGTGTCCAGCTGTGGTCCACAATCTTTCAGAGATTCTCTGAAAAGTT	5425
Db	102063	CACCTCTGTGTGCTCTCATTTCCAGCCGTGTCCAAATTTCTCTGTGAAAAGATTACCGGG	102122
Oy	5426	AACCAAGTGAAGATGTTTCCCTGTAGACAGCAATTCAGATTCT--CCGGAAGTCAGGC	5484

Db	102123	1GAGAAAGGCCCCAGATTTCCTCCCTGTAGA CAGAGATCATGATTTTCCCAAGAACGAGAC	102182
QY	5485	TTCCAGCCCTCTCTTTTCTCTGTGCCCAAGCTGCCCGGACACTTTAGCAAACTTCAGGACACCTT	5544
Db	102183	TTTCAGACACCCGCGC--CCTCTGCCCCAGCAACCTGGCACTTTAGCAAACTTCAGCAACCTT	102240
QY	5545	TACCCCACTAGACC-----	5559
Db	102241	TCCCCCACATAGACCAAGTCTTTCAGAGAGCCTCCTTCAGATGACTTCGAGTTTCTGCAA	102300
QY	5560	-----TCTGACAGAGAAAGCAGGCACTTTACATGAGTCTGTGGGAGAA	5603
Db	102301	GGAAGAGGAACTCTTTGTGTGATGGGGAAAGCAGGCACTTTACACGGAGTCTTGACGGGAGAG	102360
QY	5604	GCCATAGGCTACCGGTGTAAAGAGGCGAGGAAAGTGTGTGTGTAAGAAAGTCAGAACTTCA	5663
Db	102361	GCCATAGGCTATGGCATAGACAGAGCGCAAGGTGTGTGAAT-----TGACATTTCA	102411
QY	5664	CATAGAAAGCTTAGGCCCAACCAAGCAAAATGACACAGATCCCTCCCTATCCTCCCCCAATAGA	5723
Db	102412	CACAGAAAG-CTAGACACACACAGGAAT---GACATATCCCTCTATCTCCCCCAATAGA	102466
QY	5724	GTTTGAAG 5730	
Db	102467	GTTTGAAG 102473	

RESULT	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
5	AC119293/c	AC119293	210784 bp	DNA	linear	HTG_19-NOV-2002	
		Rattus norvegicus clone CH230-272D8, WORKING DRAFT SEQUENCE.					
		AC119293					
		AC119293.4					
		GI:25072533					
		HTG; HTGS_PHRASE2; HTGS_DRAFT; HTGS_FULLTOP.					
		Rattus norvegicus (Norway rat)					
		Rattus norvegicus					

REFERENCE AUTHORS

1 (bases 1 to 210784)

Muzny, D., Marrie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Aisbrooms, S., Amin, A., Anguiano, D., Anylebsch, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnesread, M., Benahmed, F., Bissolo, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cessari, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy, Carroll, L., De Anda, C., Dedertch, D., Delgado, O., Denson, S., Desarm, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Schoch, S., Dunn, A., Dudbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, P., Frazer, C.M., Gabrieli, A., Gante, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, M., Guetara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Haves, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huyl, S., Hume, J., Idlebird, D., Jackson, J., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Lervan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, L., Lorenshuwa, L., Lounsgesed, H., Lozado, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindradevi, M., Mamoud, M., Malloy, K., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Manwiny, S., McLeod, M. P., McNeill, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemeyor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Nair, L., Nankarva, S., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokemeloh, O., Okunonu, G., Olarunpusojun, A., Pal, S., Parks, K., Pasternack, S., Paul, H., Perez, A., Perez, L., Pfannkoch, C.,

Flopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L.,
 Puzo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R.,
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,
 Rives, C., Rodey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J.,
 Sanders, W., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,
 Shetty, J., Sivartsbeyn, A., Sison, I., Sitter, C.D., Snajls, D.,
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J.,
 Steime, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C.,
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K.,
 Valas, R., Vera, V., Villasa, D., Waldron, L., Walker, B., Wang, J.,
 Wang, O., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,
 Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K.,
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von
 Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O.,
 Weinstock, G. and Gibbs, R.A.
 Direct Submission
 Unpublished
 2 (bases 1 to 210784)
 Worley, K.C.
 Direct Submission
 Submitted (26-Apr-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 210784)
 Rat Genome Sequencing Consortium.
 Direct Submission
 Submitted (19-Nov-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Nov 19, 2002 this sequence version replaced gi:23602929.
 The sequence in this assembly is a combination of BAC based reads
 and whole genome shotgun sequencing reads assembled using Atlas
 (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described
 in the feature table below represents a scaffold in the Atlas
 assembly (a 'contig-scaffold'). Within each contig-scaffold,
 individual sequence contigs are ordered and oriented, and separated
 by sized gaps filled with Ns to the estimated size. The sequence
 may extend beyond the ends of the clone and there may be sequence
 contigs within a contig-scaffold that consist entirely of whole
 genome shotgun sequence reads. Both end sequences and whole genome
 shotgun sequence only contigs will be indicated in the feature
 table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GRM
 Center clone name: CH230-272DB
 ----- Summary Statistics
 Assembly program: Phrap, version 0.990329
 Consensus quality: 180229 bases at least Q40
 Consensus quality: 182328 bases at least Q30
 Consensus quality: 183724 bases at least Q20
 Estimated insert size: 184593; sum-of-contigs estimation
 Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 1 210784: contig of 210784 bp in length.
 Location/Qualifiers
 1..210784

```

/misc_feature
    /organism="Rattus norvegicus"
    /mol_type="genomic DNA"
    /db_xref="taxon:10116"
    /clone="CH230-272D8"
    1..1742
    /note="wgs_end_extension
clone_end:Sp6"
18139..19359
    /note="wgs_end_extension
clone_end:Sp6"
21332..22032
    /note="clone_boundary
clone_end:Sp6
site:
end_sequence:BZ243354"
200775..201644
    /note="clone_boundary
clone_end:T7
site:
end_sequence:BZ243352"
209704..210784
    /note="wgs_end_extension
clone_end:T7"

```

Query Match	40.9%;	Score 2343;	DB 2;	Length 210784;
Best Local Similarity	71.2%;	Pred. No. 0;		
Matches 4321;	Conservative	0;	Mismatches 795;	Indels 951; Gaps 55;
QY	1	GGATCCCTGCAGAGGTGCACACAAGGGTCTCCACACCACAGTGGCCCTAGTCTCAATTTCAGT	60	
Db	153822	GGATCCCTGCAGAGGTGCACACAAGGGTCTCCACACTACAGAGTGGCCCTATCTCAACTTCAGT	1537	
QY	61	TTCCATGCGCTTGTCTCACAATGCTGGCCCTCCGACGCTAATTGGACTTTGTTTAT	120	
Db	153762	TTCCATGCTGCTGCTCCCTACCAATGTTGGCCCTCCGACGCTAATTGTTAC-TTGGTTTAT	1537	
QY	121	TTCAAAAAGGGCCTGAATAGAGAGTAGATCTTGTCTACCTCCAGCTCTAAGGGTGGCCGTA	180	
Db	153703	TTCAAAAAGGGCCTGAATAGAGAGTAGATCTTGTCTACCTCCAGCTCTAAGAGTGGCTGTA	1536	
QY	181	AGCCCTCAGACTTGGAGCCCTTTGTCACAGCCCTTTAGTGTGAAGCAGATPAAAGCAATTT	240	
Db	153643	AGCCCTCAGACTTGGAGCCCTTTGTAAGACAGCACCCTTAGGCAAGAAACAAATPAAAGCAATTT	1535	
QY	241	TTCTTAAGCCAAATCTCGCTCTACACTTCTTCTGACCTCGGTCCCT-GGGCTC	299	
Db	153883	TTCTTAAGCCAAATCTCTGCTCTTAACCTTTCTTACTGACCTCGGTCCCTGGGGCTC	1535	
QY	300	TAGGGTGGGGAGGTGGGGCTTGGAAAGAAAGGTGGGGAAGTGGCAAAAGCCATGCCCTA	359	
Db	153523	TAGGGTGGGGAGGTGGGACTTG---AAAGAAAGGAGGGAAGTGGCAAAAGGCCATGCCCA	1534	
QY	360	GGGGCCCTGTGAAGTTCGGAAGCTTCCCTGTACAGACTTGGCTCATTAATCTCTCCACG	419	
Db	153466	GGACTCTGTGAAGTTCGAGGCTTCCAGTACAGACTTGGCTCATTAATCTCTCCACG	1534	
QY	420	CAAACTATGCAAGAAGTATACCTCTTGTGACTTCCGACGGCCAGTACTGTACAGT	479	
Db	153406	CAAACTATGCAAGAAGTATACCTCTTGTGACTTCCGACGGCCAGTGTGCTGTCCGGT	1533	
QY	480	TGAACAGAGATTTTAGAGAAAGCTCTGAACTCACTGAACTCTGAGCTCATCCACAAGC	539	
Db	153346	TGAACAGAGATTTTAGAGAAAGCTCTTAACTCACTGAGGCTCTGAGGCCATCTAC---C	1532	
QY	540	AAGCACTAAGGTGTCACGTGCTAGTTATATCTTACGCTGATTAATATGACAGAGCTG--GGC	597	
Db	153290	ACACACTTACATGCGGTTCTTATCAAGTATCTTACACTGATTAATACGACAGCTGACAC	1532	
QY	598	CACAGAAAGTCTGGGGGTGTAGAACTGACCAAGTACTTTTCAGTGGCAAAAGGTATAGAC	657	
Db	153230	GGGAGAAAGGCTCTGGAGTGTAGAACTGCGCAAGTACTTTTCAGTCTCAGAGGTATAGAC	1531	

QY 658 CCTCAGCAGAT---GTATGATATGTCCTTATGATCCCATCCAGGAGGCTCTCTAAG 714
Db 153170 CCTTGGCAGATATGCGAGATGGGTCCCTTAGAACCCATCCAGGAGGCTCTGAGAG 153111
QY 715 GACATGGGATGAGAGATGATGATCTGTGGCATTTCCAAACAGATATCCAGATGCCCT 774
Db 153110 GACACAGGATGACAGATGATGATCAGATGGCATTCAAACAGATATCCAGATGCTCT 153051
QY 775 TGGCCCTTCCATTTAGCCAGAGAGACAGTAACTTTAGCCTATCTTTCTCCCTCCATCC 834
Db 153050 TGTCCCTTCCATTTAGCCAGAGATGAGATAGCACTTTAGCCTATCTTTCTCCCTATCT 152991
QY 835 TCCAGAGACACACCCCTGCTGTGAGATTTCAATTTCTTCTTCACTCCCTCTGTGAC 894
Db 152990 TCCAGAGACCCACCCCTGCTGTGATTTCAATTTCTT-CTTCAATTTCCCAACATGAC 152932
QY 895 TTCCATTTGCAAGGCTTTTGAACCTGTGCAAGCTGTGAGATGAGTTTGGCCCTTAGTG 954
Db 152931 TTCCATTTGCAAGGCTTTTGAACCTGTGCAAGCTGTGAGATGAGTTTGGCTTATATG 152872
QY 955 TGGCAACCATCTTCAGAGAGAGACAGACAAACAGGAGGACAGATTTTGAAGATCAGAG 1014
Db 152871 TGGCAACCATCTTCAGAGAGAGAGAGAGAGAGAGAGAGATTTTGTATAGG-AGGAGAC 152813
QY 1015 ACTAATCACTGTGCGGCTGTGGGGGTAGAAAAAGAGTGATGATGCCCTCCAGCTAAG 1074
Db 152812 ACTAATCACTGTGCGGCTGTGGGGGTAGAAAAAGAGTGATGATGCCCTCCAGCTAAG 152761
QY 1075 CCAAGCTAGTCCCGAGATCTGTGCAAGCTGTGGGCTGTGGGGGTAGCTTATAGAGATG 1134
Db 152760 CCAAGCTAGTCTTGTGGATTT---CACAGCTGGGCTGTGCT-GAGGTAGGTTTATAGATG 152706
QY 1135 TGGGCTGAAAGACATTTGGATTTGGAAGACATCTTTTGAATCTCCCTCAACCCCACT 1194
Db 152705 TGGGCTGAAAGACATTTGGATTTGGAAGACATCTTTTGAATCTTGTGCTGTGGGCT 152646
QY 1195 ACAGACACATCTGTGTGTGGCAGATCTGTGTTAAAGCCCTGTGTGTTCTGACCATG 1254
Db 152645 AAATCTCTCCCTCAACCCACCCCATATAGAC-----ACGCTTGTGATGAGCCAGCTG 152596
QY 1255 AGCTAGGCAACAGAGATGGGCTGTGCTGAGATGAGAGATTTGTTTCAATATAGCA 1314
Db 152595 AGTTAGGCAACTGAGATGGGCTGTGTG-AGAGAGAGAG---GGTTGCAATATAGCA 152541
QY 1315 AAACAGAGGAG 1374
Db 152540 AAACAGAGGAG 152484
QY 1375 ATCAGATGAGATGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1434
Db 152483 ATCAGATGAGATGAG 152424
QY 1435 AGGAG 1443
Db 152423 AGGAG 152364
QY 1444 TTTGAGGAG 1484
Db 152363 CCGAGCTTACATCCCTGGATCTGTGATTTGCACTAGAGAGAGAGAGAGAGAGAGAGAG 152304
QY 1485 CCCAGCTCTCTCTTTATACCTCATCCC----- 1513
Db 152303 GTATATGAGCAATTTTCAATTTCAATTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 152244
QY 1514 ----- 1513
Db 152243 TACATTTGTTACTCTATACAGATATGCTTATGAGAGAGAGAGAGAGAGAGAGAGAGAG 152184
QY 1514 ----- 1513
Db 152183 ACAAGTCTTTTATGAAATCAGTGTGATTTAATCAAGCTGTGATGAGAGAGAGAGAG 152124
QY 1514 -----GTCTCCCAATTAAAGCCCACTCTTCTTCTAGATCAGAC 1551

Db 152123 CAG 152064
QY 1552 CTGAGCTGAG 1511
Db 152063 CTCTTC-CTGAG 152005
QY 1612 AAGTCCAG 1671
Db 152004 CAGTGAAG 151945
QY 1672 GTGGCAGAGCTGTG--CACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1729
Db 151944 GCAAGTACAG 151885
QY 1730 TGATGCTCAG 1789
Db 151884 TGATGCTCAG 151856
QY 1790 GCCTGAG 1849
Db 151855 ---GGAGCTGAG 151800
QY 1850 ACCTTTAGATGTATATCATGAG 1909
Db 151799 ACCTTTAGATGTATATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 151744
QY 1910 CCGAGGAG 1969
Db 151743 ---AATTAAG 151704
QY 1970 TGAAGGAG 2029
Db 151703 ---TCCGATATGAG 151656
QY 2030 GGGAG 2089
Db 151655 GGGAG 151596
QY 2090 GTCTTCTCACTGTGAG 2148
Db 151595 GTGCTCTCACTGTGAG 151536
QY 2149 AGAAGTCAAG 2208
Db 151535 AGAAGTCAAG 151477
QY 2209 GGGAG 2268
Db 151476 GGGAG 151417
QY 2269 CTGAGTTAACTCTTCTTGTGATCATCTTCAATAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2328
Db 151416 CTGATTTAACTCTTCTTGT---ATCTTCAATAGAGAGAGAGAGAGAGAGAGAGAGAG 151360
QY 2329 CATCCCATGAG 2388
Db 151359 CAT-----CCCCACTATCTATACCAATTTTGGCTGTGATGAGAGAGAGAGAGAGAG 151310
QY 2389 CTGAG 2448
Db 151309 CTGAG 151250
QY 2449 GTACCTTTGAG 2508
Db 151249 GTGCTTATCATCTCTCAAG 151190
QY 2509 AACCTTGAAGCTTGAAG 2568
Db 151189 AACCTTGAAGCTTGAAG 151130
QY 2569 TGGCCTTTATCCCAAG 2628

Db 151129 TGGCTTCATCTCGGAGAGCAACATTGGCAGAGGACCAAGACAGATGGGAAACGCAAG 151070
 QY 2629 TTGAATCAGTCCCTTCAAGAAATACGTCATGCAAGACCTTAAGCCCTTGAAGAGGGGT 2688
 Db 151069 TTGAAACAGGTCCCTTCAAGAAATACCGCGTCAAGATCTTAAGCCCTGAGAGGACTGT 151010
 QY 2689 ATGCTCTGCGCCCAACCCACCATPAAGGGAGTGAATCTATCTTGGGGGCTGGGCACTTGG 2748
 Db 151009 ATGCTCTGCAACCCACCCACCATPAAGGAGTGAATATCTTAAGGGGCTGGGCACTTGG 150950
 QY 2749 GGGAGACACCACTTAATACTAGAGTGTGAGCCCGAAGAAAATGACCCCTGTGTCTGC 2808
 Db 150949 GAGAGACACCACTCAACGAGAGTGTGATCTCAGAAACCTGACCGCTGTGTCTCTAC 150890
 QY 2809 CCACCTTCCACACTTGAAGTAAAT--GAGAGTGAACATAGATAGGAGTGGAGCTGGT 2866
 Db 150889 CCACCTTCCACACTTGAAGTAAATTTGAGAGAGTGAACATAGAT--GGGAGAGACGTGGG 150831
 QY 2867 AGCAGGAGAGTGTCTGGGAGTGAAGGGGTGAGGGGAAAGCCAGAGCAAGGGAGTCTGG 2926
 Db 150830 AGCAGGAGAGCTTCTGGGAGTGTGGGGGTGGA--GGAAAGTCAAGACAGGGAGTCTGG 150772
 QY 2927 CTTTGTCTCTGAACAACAATGTCTACTTATTAACAAGCATGACCTGTAAAGACCA 2986
 Db 150771 CTTTGTCTCTGAACAAGTGTCTACTTATTAACAATGACATGACCCGCTAAAGACCA 150712
 QY 2987 ACATCTACACCTCTGAAAAGACAGCCCTTGAAGACAGGGGTGTCTCTGAGCCTTG 3046
 Db 150711 ACACCTACACCTCTGAAAAGACAGCTCTTGAAGACAGGGGTGTCTCTGAGCCTTG 150652
 QY 3047 GGTGCTTGAATGTGACCAAAAGAGGAGCATGATGTGAGTATPAAGGCCCCAGAGAGCTTA 3106
 Db 150651 GGTACTTGAATGTGACCAAAAGAGGAGCATGATGTGAGTATPAAGGCCCCAGAGAGCTTA 150592
 QY 3107 GAGAAGGACACTTGGAGAGGGGTCACTGTGACAGAGCCCTTATCATGATGAATCTGAGCCT 3166
 Db 150591 GAGAAGGACACTGGGAGGGGTCACTGTGACAGAGCCCTTATCATGATGAATCTGAGACA 150532
 QY 3167 GGGGCCAATGTGTAAATCTTGGGCTGTGACAGGCAATTCAAAGACAGCATCTGCATCTC 3226
 Db 150531 GGGGCCAATGTGTAAATCTTGGGCTGTGACAGGCAATTCAGAGTACCATCTGCATCTC 150472
 QY 3227 TGGCAGCTGGGGAGGGGAGGAGCAACCCGCCCTTAATACCTTCTCCCTGAGCCC 3286
 Db 150471 TGGCAGCTGGAGAGGGGAGGAGCAACCCGCCCTTAATACCTTCTCCCTGAGCAC 150412
 QY 3287 CAGGATTTAACACTCTGAGCTTCCCTTCCACCTTCCATCAGAGTGAAGGGTTCAG 3346
 Db 150411 CAGGATTTAACACTCTGAGCTTCCCTTCCACCTTCCATCAGAGTGAAGGGTTCAG 150352
 QY 3347 AGGAGGGTAAACCTACATGTCCAAACATCATGTGCAAGATATATGATCATATGT 3406
 Db 150351 AGGAGGGTAAACCTAGAT-----AATGTGTGATGATATATGCACTAGATGT 150300
 QY 3407 GTAGAGGCAAGAAAGAAATCTGCA-----GGCTTAATCTGGTTAATGTGTA---GT 3457
 Db 150299 GTAGAGGCAAGAAAGAAATCTGAGTTAATGTGTAAGTCTGTGTGTGTCAGATGTG 150240
 QY 3458 CTGTGTGATGTGTGTGTGTGTGCTGATGAAACCGGAGTGGCTGTGAGCTGTTCAGTTCT 3517
 Db 150239 ATGTGAATGATGTGTGTGTGTGTGATCTGAAGAGTGGGATGGAATGTGACGTTCAGTTCT 150180
 QY 3518 GTGCGTGAAGTTAACAGACTGCAAGTTTGTGTGTAATTTGCCAAGGCAAGGTGGGTGA 3577
 Db 150179 GTGTGTGACATTAACAGACTGCAAGTTTGTGTGT-----AAGTCACTGAG 150135
 QY 3578 TCCCTTCAATGTGTTAAAGATTTGATGATGAGCTGCAATTCAGAACCATGGAATAA 3637
 Db 150134 ACCCTCCCATGTGTTAAAGAAATTTGATGATGAGCTGCAATTCAGAACCATGGAATAA 150075
 QY 3638 GAATGACACTCTATATGTGTCTCTAAGCTAAGTAGCAAGTCTTTGAGAGCACCTGT 3697
 Db 150074 GAATGATGACTTATATATGTCTCTAAGCCAAAGTAGCAAGTCTTTGAGAGCACCTGT 150015

QY 3698 CTTAGAGTGTGGGCAACAGAGCTACAGACATCTGTATCTGTACAGATTAAGAGAGAGA 3757
 Db 150014 CTTAGAGATGTGGCAACAGAGCAATTAACATATCTGTACAGATTA--AAGAGAGAG 149957
 QY 3758 GGGGTGTGAATTTCTTACTATCAAGAGGAACTGATCTGTGACCTGCATAAGTGTAT 3817
 Db 149956 TGAAGTATGAATTTCTTACTATCAAGAGGAACTGATCTGTGACCTGCATAAGTGTAT 149897
 QY 3818 GCTTCTCTTGAACATCATGACTTTGTCTGTGGGAGCCAGCATGTGTGAATCTCAGTCT 3877
 Db 149896 GCTTCTCTTGAACATCATGACTTTGTCTGTGGGAGCCAGCATGTGTGAATCTCAGTCT 149837
 QY 3878 GAGAGATGAAGAGGTCTCCCTCAGCTGAAGCTATGCAATTAAGCAGGGTTGAA--GGG 3935
 Db 149836 --GAGAGTGAAGAGGACCCCTTACGCTTAAGCTGTGCAATTAAGGAGTGAAGAG 149779
 QY 3936 GGAAGAGAGAGCTGTGAGTGGAGCTGTGTGTGGAAGGAGGAGCAGATTTAAGCCT 3995
 Db 149778 GAAAGGGAGAGCTGTGAATGGAGCTGTGTGTGTGAGACAGGGGACAAATATTAAGCCC 149719
 QY 3996 GGAAGAGAGAGTGAACCTTAACCACT--TGTCACTCAACCTTCAGATTTAAATTAAC 4054
 Db 149718 GTAAGAGAGAGTGAACCTTAACCACTGTGTCACTCAGCTTCAGATTTAAATTAAC 149659
 QY 4055 AGGTAAAGGCC---TGGTGAAGGAGGTGTGTGTGAAGCCTCTCTCTCTCTGCAATG 4111
 Db 149658 AGTAAAGGCAATGTGGGTGAAGGAGGTGTGTGAAGCCTCTCTCTCTCTCTATCTG 149599
 QY 4112 CCTGAGGCCCTTTGGGAGAGGAGAAATGTGCCAAGCACTAATAAAGGCCATGAGCCA 4171
 Db 149598 CCTATCGGCCCTTTGGGAGAGGAGAAATGTGCCAAGCACTAATAAAGGCCCTGAGCCA 149539
 QY 4172 GAGGGGCAAGGCAACAGACTTTTATGAGGCAAACTTTGGGGCCGTAGTATGATGA 4231
 Db 149538 GAGGGGCAAGGCAACAGACTTTTATGAGGCAAACTTATGAGGCAAACTTATGAGGCA 149496
 QY 4232 CAAGAATCTGCAATCAATCACTCTCTCTCTTCTAAGCGACAGAGAGGAATCGAATTAC 4291
 Db 149495 -----TGCTGTCTCTCTCTCTCTCTCACTTCCAGAGCCAAAGTCAAGAGGAG 149451
 QY 4292 CACTCCATCAGATGTAGAGAAAGTGAAGTGAAGTTTACCTCCCTATCAGTATGA 4351
 Db 149450 GAGCCAGAGCAGAGAGGATGGAGGAG----- 149423
 QY 4352 GAGAAAGTGAAGTGAAGTGAAGTTTACCACTCCCTATCAGATGAAGAGAAAGTGA 4411
 Db 149422 ----- 149423
 QY 4412 GTTTACACTCCCTATCAGATGTAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGA 4471
 Db 149422 ----- 149423
 QY 4472 GTGATGAAGAAAGTGAAGTGAAGTTTACCACTCCCTATCAGATGTAGAGAAAGTGA 4531
 Db 149422 ----- 149423
 QY 4532 AGTCAGTTTACACTCCCTATCAGATGTAGAGAAAGTGAAGTGAAGTGAAGTGAAG 4591
 Db 149422 -----GATCCAG 149415
 QY 4592 CAGAGACTCCAAATTTAGGACGACGCAATATGAGATGATTAAGGGCTGAGAGCAC 4651
 Db 149414 CAGATGACTCCAAATTTAGGACGACGCAATATGAGATGATTAAGGGCTGAGAGCGC 149355
 QY 4652 TGAGAGCTGTCA---GAGATTTCTCAACCCAGGTAAAGGAGATTTGAGGTGGGGGT 4707
 Db 149354 TGAGAGCTGTCAACCCAGATTTTCTCATTCAGGTTAAGAGAGATTTAGCGTGGGGGT 149295
 QY 4708 CTTCAACCAACAGACTCTCCCACTAGAGGAAACCTGCTTCTCTGAAAGTGGGT 4767
 Db 149294 CTTCAACCCAGCAACAGACTCTCCCACTAGAGGAAAGTGTCTTCTCTGAAAGTGGGT 149235


```

QY 4768 TCAGCCCGGTGAGAGATCTGACAGAGGTGGCTTCCACAGCCCTGGAGAGTTCTCAGTGGC 4827
Db 149234 T-----CTCCCAACCGGCTGGGAGTTCTCGGTGGC 149203
QY 4828 AGGAGTTTCCACAGAGAAACATGATGCCCTTCCCTTACGCTGTCTTCTCATCTTCC 4887
Db 149202 AGGATGTTTC-----TACGAGATGCCCTTCCCTTCCACATGCCCTCTCCCTC----- 149156
QY 4888 TCTGGGGAATGATCTTCCCGCTGGTTATCTGGATCTTGGCTCTTCTGCTGACAGAAATTTG 4947
Db 149155 -----CTTGTTCTTAAATTAATTTGGCTTTAGTTGTTGAGAAAGATTTG 149112
QY 4948 CCTGTGCTGTCACATCTTCATCTTCTCTACTGTCTCCGCTGCTGCTGCTTCCCTTCCG 5007
Db 149111 CCGGAGCTGTACTCATCTGTCTCTCTCTC-----TGCCTTCCCTTCTTGAG 149059
QY 5008 TGTCTCTCTTTCACCCATTTTCTCATCTTCACTTTTCTCCCTTCTCATTTTGTATTCAT 5067
Db 149058 TGTCTCTCTTTCACAGTGTCTCTCATCTCAGCTGCCCTCCCTCCCTTCATTTTAA- 149000
QY 5068 CCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5127
Db 148999 ----- 149000
QY 5128 CCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5187
Db 148999 -----TCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 148959
QY 5188 TGGGTTCC--CACCTTAATTAATTAATCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5245
Db 148958 CAGGCTTCAATACAGTCAAGTAAAGCAATCTCCAGAGACCAAGCTTTAATCT-CTG 148900
QY 5246 GGTGCTCTCTTACCTTCTCTCAACCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5305
Db 148899 GGTGCTCTTACCTTCTCTCACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 148840
QY 5306 AGATGTCTCTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5365
Db 148839 -GGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 148781
QY 5366 CGGCATCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5425
Db 148780 CACTCTGTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 148721
QY 5426 AACCAAGTAAATGTTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5484
Db 148720 TGAAGATGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 148661
QY 5485 TTCCAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5544
Db 148660 TTCCAGCAACCGC--CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 148603
QY 5545 TACCCCACTATGAC----- 5559
Db 148602 TCCCTCCATATGACCAAGCTTCTGAGAGAGCTCTCTCTCTCTCTCTCTCTCTCT 148543
QY 5560 -----TCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5603
Db 148542 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 148483
QY 5604 GCAATAGGCTTACGCTGTAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5663
Db 148482 GCAATAGGCTTACGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 148432
QY 5664 CATGAAAGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 5723
Db 148431 CACAGAAAG-CTAAGACACACAGAGAGAT---GACATATCTCTCTCTCTCTCTCTCT 148377
QY 5724 GTTTGAG 5730
Db 148376 GTTTAAG 148370

```

```

RESULT 6
AY191158
LOCUS
DEFINITION
Rattus norvegicus beta myosin heavy chain gene, partial cds;
beta-alpha myosin heavy chain intergenic spacer region, complete
sequence; and alpha myosin heavy chain gene, promoter region.
ACCESSION
AY191158
VERSION
AY191158.1 GI:32493126
SOURCE
Rattus norvegicus (Norway rat)
ORGANISM
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 5487)
Haddad, F., Bodel, P.W., Qin, A.X., Giger, J.M. and Baldwin, K.M.
Role of Antisense RNA in Coordinating Cardiac Myosin Heavy Chain
Gene Switching
J. Biol. Chem. 278 (39), 37132-37138 (2003)
JOURNAL
PUBMED
2 (bases 1 to 5487)
Haddad, F., Qin, A.X. and Baldwin, K.M.
Direct Substitution
Submitted (04-DEC-2002) Physiology and Biophysics, University of
California Irvine, Medical Science 1, D360, Irvine, CA 92697, USA
FEATURES
source
1..5487
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/chromosome="15"
/map="15p13"
join(<1..136,703..824)
/product="beta myosin heavy chain"
join(<1..136,703..720)
/codon_start=2
/product="beta myosin heavy chain"
/protein_id="AA034581.1"
/db_xref="GI:32493127"
/translation="EEQANTNLSKPRKQVHLEDAERADIAESQVNLKRAKRDIGA
KGLNER"
825..5024
/note="beta-alpha myosin heavy chain intergenic spacer
region"
misc_feature
complement(2868)
/note="antisense transcription start site"
misc_feature
complement(2905)
/note="antisense transcription start site"
gene
5025..55487
/gene="alpha myosin heavy chain"
5025..5030
/gene="alpha myosin heavy chain"
ORIGIN
Query Match 40.5%; Score 2320.2; DB 10; Length 5487;
Best Local Similarity 77.5%; Pred. No. 0;
Matches 3475; Conservative 0; Mismatches 583; Indels 425; Gaps 38;
QY 1 GATTCCTGCAAGTCAACAAGGCTCTCCACCAAGAGGCTGTGTCATTTCACT 60
Db 576 GATTCCTCAAGGTCAACAAGGCTCTCCACCAAGAGGCTGTGTCATTTCACT 635
QY 61 TTCCATGCTTGTTCACAAATGCTGACCTCCCAAGAGCTAATTTGACTTTGTTAT 120
Db 636 TTCCATGCTGTGCTCTCAAAATGTTGGCTCCCAAGAGCTAATTTTGAC-TTGTTTAT 694
QY 121 TTCAAAAGGCTGAATGAGAGTGAATCTTGTGCTAACCAAGCTTAAGGCTGCCGTGA 180
Db 695 TTCAAAAGGCTGAATGAGAGTGAATCTTGTGCTAACCAAGCTTAAGGCTGCTGA 754
QY 181 AGCCCTCAAGCTTGAGCTTTTGCAACAGCCCTTTTGTGGAAGCAATTAAGCAATT 240

```


Db 755 AGCCCTGAGACTGAGACCTTTGAAAAGACCTTAGAGCAAAACAAATTAAGCAATT 814
QY 241 TCCCTTAAGCAAAATTCCTGCTTCTGACTCTTCTCTGACCTGGTCCCT-GGGCTC 299
Db 815 TCCCTTAAGCAAAATTCCTGCTTCTGACTCTTCTCTGACCTGGTCCCTGGGCTC 874
QY 300 TAGGGTGGGAGGTGGGCTTTGGAGAAGAGGTGGGAAAGTGGCAAAAGCCGATCCCTA 359
Db 875 TAGGGTGGGAGGTGGGACTTGAAGAGGA---GGGAAAGTGGCAAAAGGCCACTCCCA 931
QY 360 GGGCCCTGTGAAGTTCGAGCCCTTCCCTGTACAGACTGGCTCATAGATCTCTCTCAGC 419
Db 932 GGACTGTGTGAAGTTCGAGCCCTTCCAGTACAGTCTGGCTCATAGATCTCTCTCAGC 991
QY 420 CAAACATAGAAAGATGATCTCTCTTGTGACTTCCCAAGGCCAAGTACCTGTAGGT 479
Db 992 CAAACACAGAAAGATGATCTCTCTTGTGACTTCCCAAGGCCAAGTACCTGTAGGT 1051
QY 480 TGAACAGATTTAGAGAGCCTCTGAACTCAGCTGAACTTGAAGCTCATCCAGCAAG 539
Db 1052 TGAACAGATTTAGAGAGCCTCTAATCTCAGCTGAACTTGAAGCTCATCTAC---C 1107
QY 540 AAGCACTTAGTCCACTGTAGTATCTTACGCTGATTAATATATGACAGCTG--GGC 597
Db 1108 ACACACTTACATGCGGTTCTAATCAGTATCTTACACTGATTAATATGACAGCTGAC 1167
QY 598 CACAGAAATCTGGGGTGTAGAACTGACCAAGTACTTTCACTGGCAAAAGTATGACC 657
Db 1168 GGGAGAGAGGCTGAGATGAGAACTGGCCAGTGACTTTTCAGCTGACAGAGTATGACC 1227
QY 658 CCCTCAGACAGAT---GTAGTAATGTCCCTTATAGTCCATCCAGGAGGCTCTAAGAG 714
Db 1228 CCCTTGGCAGATATGAGCAGATGGGTCCCTTAGAACCATCCAGGAGGCTCTGAGAG 1287
QY 715 GACATGGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 774
Db 1288 GACACAGATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1347
QY 775 TGCCCTTCCACTTATGAGCAGAGAGACATTAACCTTATGCTTCTTCTCTCCATCC 834
Db 1348 TGTCCTTCCACTTATGAGCAGAGAGATGAGCAACCTTATGCTTCTTCTCCATCT 1407
QY 835 TCCAGAGACACACCCCTGTGTGAGATTAATTTCTTCTCAGTCCCTCTGAGAC 894
Db 1408 TCCAGAGACCAACCCCGGTCTAGTATTAATTTCTTCTCAGTCCCTCTGAGAC 1466
QY 895 TTCCATTTGCAAGGCTTTTGAACCTCTGACAGCTCTGAGATAGATTTGGCCCTAGT 954
Db 1467 TTCCATTTGCAAGGCTTTTGAACCTCTGACAGCTCTGAGATAGATTTGGCTCTATAT 1526
QY 955 TGGCAAGCATCTGAG 1014
Db 1527 TGGCAAGCATCTGAG 1585
QY 1015 ACTTAATCACTGGGCGGCTGGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074
Db 1586 ACTTAATCACTGGGCGGCTGGGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1637
QY 1075 CCAAGTATGCTCCGAGATATCTGTGCAAGCTGGGCTGTGGGAGTGAAGAGAGAG 1134
Db 1638 CAAAGCAGATCTTGGGTATTT---CACAGCTGGGCTGT---GAGGTAGTGAAGAG 1692
QY 1135 TGGGTCTGAAGAGAGATGGAGTGAAGAGATCTTTTGAAGTCTCCCTCAACCCACT 1194
Db 1693 TGGGTCTGAAGAGAGATGGAGTGAAGAGATCTTTTGAAGTCTCCCTCAACCCACT 1752
QY 1195 ACAGACACATCTGTGTGGCAGAGCTCTGTCAACAGCCCTGTGTGTCTGACCACTG 1254
Db 1753 AAATCTCCCTCAACCCCACTTACAGAGCTTGTGACAGCTGAG---ACTG 1802
QY 1255 AGCTAGGCAACAG 1314
Db 1803 AGTTAGGCAACAG 1857

QY 1315 AAACAGAGGGGAG 1374
Db 1858 AAACAGAGGGTGGAG 1914
QY 1375 ATCAGATGAGTCAAG 1434
Db 1915 ATCAGATGAGTCAAG 1974
QY 1435 AGGAG 1443
Db 1975 AGGAG 2034
QY 1444 TTTGGGCAATTAAG 1484
Db 2035 CCGACTTATACCTCGGTGAGTGTCTATTTGCACTCAGAAATACAGACAGAGAACT 2094
QY 1485 CCAAGTCTCTCTTATTAACCTCAATCC--- 1513
Db 2095 GTATATGGCATTTTCACTTCTCAATCTGAAGCTTGAACACTTGAATTTCTAGTT 2154
QY 1514 ----- 1513
Db 2155 TACATTTGTACTCTATACAGATATGTCTTCATGAGACAGTGTGTGATCTCCTGAAT 2214
QY 1514 ----- 1513
Db 2215 ACAAGTCTTTTTTTTATGAATCAGTGTGATTAATCAAGTCTGCATAGAGAGCCA 2274
QY 1514 -----GTCTCCCAATTAAGCCCACTCTTCTCTAGATCAGAC 1551
Db 2275 CAGAAAGAGCCAGAGTCCCAAGTCTCTCTTATAGCCCTCCCTCCTTAAGCCATTC 2334
QY 1552 CTGAGCTGAG 1611
Db 2335 CTCTTC-CTGAAACAGACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATCA 2393
QY 1612 AAGTCCAGGAGCAACCTTAAGAGCAGAGCCTTAAGAGAGAGAGAGAGAGAGAGAG 1671
Db 2394 CAGTGAAGAGAGATGTGAGACAGAGCCTTAAGAGAGAGAGAGAGAGAGAGAGAGAG 2453
QY 1672 GTGGCAGGCTGTG---CACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1729
Db 2454 GAGCTAGCCCTGACACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2513
QY 1730 TGAATGCTGAATGGAG 1789
Db 2514 TGAATGCTGAATGGAG 2541
QY 1790 GCTGGAACAGAGATCTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1849
Db 2542 ---CGGAGCTGAGACTGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2598
QY 1850 ACCTTTGAATGTAATTCATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1909
Db 2599 ACCTTTGAATGTAATTCATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2658
QY 1910 CCGAGGAG 1969
Db 2659 -----AGACAG 2695
QY 1970 TGAAGCAG 2029
Db 2696 -----CGATATGAG 2742
QY 2030 GGAAGCAG 2089
Db 2743 GGAAGCAG 2802
QY 2090 GTCTTTCTCACTGTGGCAG 2148
Db 2803 GTGCTCTCACTGTGGCAG 2862

QY	2145	AGACTGCAAGGCGCTGGCTTTGGGAGGCGAGCTGGAAAGAGTAATGTGAGACGACGAGGGAGCA	2208
Db	2863	AGACTGCAAGGCGCTG3TGTGGAGGCGCGCTGGAAAGAGTAATGTGAGAG- CGG3GGAGACTA	2921
QY	2209	GGGGCGCTAGAAAGAAAGAAAGAGAGGGCCAAACCAAGGCCACACAAAGAGGGCAGAGCCGAGAA	2268
Db	2922	GGGGACCGAGGAAAGAAAGAAAGAGAGGGCAAAACAGGCCACACAAAGAGGGCAGAGCCGAGAA	2981
QY	2269	CTGAGTTAACTCCTTCCTTGGTGCATCTTCCATAGAGAGCAGTGGAACTCTGTGACCAAC	2328
Db	2982	CTGATTTAACTCCTTCCTTGTG- -ATCTTCATAGAGGCGCAATGGAAACTCTGTGCCAC	3038
QY	2329	CATCCCCCATAGGCCCCCACTAACCCATACCAAGTTTGGCCGAGTGGCAATTCTAGTTC	2388
Db	3039	CAT-----CCCCCACTAACCTTAACCAAGTTTGGCCGAGTGGCAATTAGTGTCC	3088
QY	2389	CTGAGGACAGAGCGCTGGCGCTTGTCTCTTGGACCTGACCCCAAGCTGACCCCAATGTTCTCA	2448
Db	3089	CTGAGGACAAAGCGCTGGTTTTTGTGCGTTGACCTGACCCAGGCTGACCCCAATGTTCTCA	3148
QY	2449	GTACCTTTGAATGCCCTCAAGAGCTTGAAGAACAGGACGTGACATATTAGGCCATGGGCT	2508
Db	3149	GTGCTTTATCATGCGCTCAAGAGCTTGAAGAACAGGACGTGACATATTAGGCCATGGGCT	3208
QY	2509	AACCTTGGAGTTTGACACAGAGCGTCAATGACCTCCAGGGACACAGCTGACAGACAG	2568
Db	3209	AACCTTGGAGTTTGACACAGAGCGTCAATGACCTCCAGGGACACAGCTGACAGACAG	3268
QY	2569	TGCGCTTTATCCCAAGAGCAACCATTTGGCATATAGTGGCTGACAAATGGGAAATGCAGG	2628
Db	3269	TGCGCTTTATCTCTGGAGAGCAACCATTTGGCATAGGACACAGCAGATGGGAAAGCAAG	3328
QY	2629	TTGAATCAGTCCCTTCAGAAATACCTGCATGCAAGACTTAAGACCCCTGAGAGAGGGT	2688
Db	3329	TTGAATCAGTCCCTTCAGAAATACCGGTCATCAAGATCTAAGACCCCTGAGAGAGACTGT	3388
QY	2689	ATGCTCCGCGCCCCACCCACCATAGAGGGAATGAACTATCCTAAGGGGCTGGCCACTTGC	2748
Db	3389	ATGCTCCGCGCCCCACCCACCATAGAGGGAATGAACTATCCTAAGGGGCTGGCCAACTTGC	3448
QY	2749	GGGAGACACCACTTACTAGAGAGTCTAGCGCCAGAAAACTGACCGCGCTGTCTCTGC	2808
Db	3449	GAGAGACACCACTATCACCGAGAGTCTTACTCAGAAAACTGACCGCTGTGTCTCTAC	3508
QY	2809	CCACCTTCACACTCTAGAGCTAAT- -GAGAGGTGACAGTGAATAGGGTGGAGCTGGT	2868
Db	3509	CCACCTTCACACTCTAGAGCTAATATTTAGAGAGGTGACAGTGAAT- GGGGGAGAGCTGG	3568
QY	2867	AGCAGGAGAGATGTTCTCTGGGTGTAGGGGTGTAGGGGAAACCGAGAGAGGGAGTCTGG	2926
Db	3568	AGCAGGAGAGATGTTCTCTGGGTGTGGGGGTGGA- GGGAAATCTAGAGAGGGAGTCTGG	3626
QY	2927	CTTTGTCTCCGGAACACAAATGCTACTAGTTATTAACAGGAGATGACCTGCTAAAGACCA	2986
Db	3627	CTTTGTCTCTGGAACACAGTGTCTACTAGTTATTAACAGATGACCCGCTAAAGACCA	3686
QY	2987	ACATCTAGACCTCTGAAAGACAGACAGCCCTGAGGAGCAGGGGTGTCTCTGAGCCTTG	3046
Db	3687	ACACTAGACCTCTGAAAGAAACAGCAATCTCGAGAGCGGGGGTGTCTCTGGGCTTG	3746
QY	3047	GGTGTCTATGTGTGACAAAGAGAGGCGATAGTGTAGTATTAAGGCCCAAGACCTGTA	3106
Db	3747	GGTCTTATGTGTGACAAAGAGAGGCGATGATGTGATTAAGGCCCAAGAGGGCTTA	3806
QY	3107	GAGAAAGGCACTTGGGAAAGGGGTGAGTGGACAGACCCCTAATCCATGGAAATCTGGAGCT	3166
Db	3807	GAGAAAGGCACTCGGAAAGGGGTGAGTGTGAGAGGCCCTAATCCATGGAAATCTGGAGCA	3866
QY	3167	GGGGCAACTGGTGTAAATCTCTGGGCGTGCAGAGCAATTCAAGCAGACACTGTACCTCT	3226
Db	3867	GGGGCAACTGGTGTAAATCTCTGGGCGTGCAGAGCAATCCGAGATAGCACTGTACCTCT	3926
QY	3227	TGGCAGCTGGGGAGGCGAGAGGCAACCCCACTTAATACCTTTCTCCTCAGCC	3286

D	3927	TGGCAGCCTGGAGAGGCGGAAAGAGCAACCCCCCACTTACACCCCTTCTCCATCAGCAC	3986
O	3287	CAGGATTTAACACCTCTGACCCTTCCCTCTCCCACTCCCATCAGAGTGAAGGCTGCAG	3346
D	3987	CAGGATTTAACACCTCTGACCCTTCCCTCTCCCACTCCCATCAGAGTGAAGGCTGCAG	4046
O	3347	AGGAGAGGATAAAAACCTCATGTCCAAACATCATGTGACAGATATATGATCAGTATGT	3406
D	4047	AGGAGAGGATAGAACCTTAAT-----AATTGTGTGATGATATATGACTAGTATGT	4098
O	3407	GTAAGAGCAAGAAAGAAATCTGCA-----GCTTAACTGAGTTATGTGTAA---GT	3457
D	4099	GTAAGAGCAAGAAAGAAATCTGAGTTAATGTGTAAAGTCTCTGTGTGTGTGATGTGC	4158
O	3458	CTGTGTGATCATGTGTGTGTGTCTGACTGAAAGCGGCATAGCGTGTGCACCTTGAGTTCT	3517
D	4159	ATGTGAATGCATGTGTGTGTCTGACTGAAAGCGGCATGAGATGTGCACCTTGAGTTCT	4218
O	3518	GTCGTGAGGTTACCAAGCTGCAGGTTGTGTGTAAATTTGCCAAGCAAGTGGGTGAA	3577
D	4219	GTAATGTGACATTACCAAGCTGCAGGTTGTGTGT-----AAGTCAGTGAAG	4263
O	3578	TCCCTTCCATGCTTTAAAGAGATGTGATATGGCTGCATCTCAAGAACATGTGAAATA	3637
D	4264	ACCTCCCATGTTTAAAGAAATTTGATGATATGCTGCATGTGCAGGAAACATTGAAGTA	4323
O	3638	GAATGACACTCATATATGTGTCTTAAGGTAAGTAGAGCAAGGTCTTTTGAAGACACTGT	3697
D	4324	GAATGTAGACTTATATATATGTCTTAAGGCAAAAGTAGAAGTCTTTGAGAGACACTGT	4383
O	3698	CTAGAGATGTGGCAACAGAGACTACAGACAGTATCTGTACAGATTAAGAGAGAGAGA	3757
D	4384	CTAGAGATGTGGCAACAGAGCAATAGACAGTATCTGTACAGATTA--AGAAGAGAGG	4441
O	3758	GGGGGTGTGAATTCCTTACTATCAAGAGGAACTGAGTGTGTGCACTGCAGAAAGTGAT	3817
D	4442	TGAGGTATATGAATTCCTTACTATCAAGAGGAACTGAGTATGTGCACTGCAGAAATGAT	4501
O	3818	GCTCTCCCTAGACATCATGACTTTGTCTCTGTGGAGGACAGCACTGTGGAACCTTAGGCT	3877
D	4502	GCCTCCCTGACATCATGACTTTGTCTCTGTGGAGGACAGCACTGTGGAACCTTAGGCT	4561
O	3878	GAGAGATGTAGAGGCTCCCTCAGCCTGAAGCTATAGCAGATAGCCAGGGTTGAAA--GGG	3935
D	4562	--GAGGTAGAGGACCCCTCAGCCTGAAGCTGTGCAGATAGCTTAGGGGTGTAAAGAGG	4619
O	3936	GGAAGGAGAGCTGTGGATGTGGAGCTTGTGTGTGGAAGCAGGGGACGATATTAAAGCT	3995
D	4620	GAGAGGGGAGAGCTGTGAATGTGGAGCTTGTGTGTGTGGAAGCAGGGGACAAATATTAAAGCC	4679
O	3996	GGAAGAGAGGTGACCTTTAACCAAGT--TGTCACTCACCTTCAGATTAAATAATGCTG	4054
D	4680	GTAAGAGAAAGTGAACCTTACCAAGTGTGTCACTCACCTTCAGATTAAATAATGCTG	4739
O	4055	AGGTAAGAGGCG---TGAGTATAGGGAGAGTGTGTGAGAGCGCTCTGTCTCTCTGCATG	4111
D	4740	AGGTAAGAGGCGCATGTGGGTATGGGAGAGTGTGTGAGAGCGGTCTGTCTCTCTCATGTG	4799
O	4112	CCCTGAGGCCCTTTTGGGAGAGAGAAATGTGCCAAGACTTAAAAAAAGCCATGGAGCCA	4171
D	4800	CCCATGGCCCTTTTGGGAGAGAGAAATGTGCCAAGACTTAAAAAAAGCCCTGGAGCCA	4859
O	4172	GAGGGGCGAGAGGACAGACCTTTTACGTGGCAAACTTTGGGGC	4214
D	4860	GAGGGGCGAGAGGACAGACCTTTTATGTGGCAAACTCTAAGGGC	4902

RESULT 7				
AC115371/c				
LOCUS	AC115371	261901 bp	DNA	linear
DEFINITION	Rattus norvegicus clone CH230-118c1, WORKING DRAFT SEQUENCE.			
ACCESSION	AC115371			

```

VERSION      AC1153371.5  GI:30521556
KEYWORDS     HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP
SOURCE       Rattus norvegicus (Norway rat)
ORGANISM     Rattus norvegicus

```

REFERENCES

(Bases 1 to 263901)

Muzny, D., Matle, E., Metzker, M., Lee, A., Abramzon, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D., Anyalibech, V., Aoyagi, A., Ayodeji, M., Bacc, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Bismilio, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, K., Cree, A., D'Souza, L., Davila, M. L., Davis, C., Davy-Carroll, L., De Anna, C., Dederich, D., Delgado, O., Denison, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotroc, M., Eugene, C., Evans, C. A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C. M., Gabisai, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamli, C., Hamilton, C., Hamilton, J., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hognes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovac, C., Kowals, C., Kraft, C. L., Lebow, H., Levay, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuwal, L., Louleaged, H., Lozada, R. J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapa, P., Martin, K., Martin, R., Martinez, E., Mawhinney, S., McLeod, M. P., McNell, T. Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Mundaasa, M., Murphy, M., Nair, L., Nankervils, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwachilemeh, O., Okunolu, G., Olarnpunagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Piopler, F., Poldrester, A., Popovic, D., Primm, E., Pu, L. L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M. A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., Sanders, M., Savery, G., Scheerer, S., Scott, G., Shateman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C. D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valae, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willison, R., Wlaczek, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Neiderhausen, A., Weiss, R., Smith, D. R., Holt, R. A., Smith, H. O., Weinstein, G., and Gibbs, R. A.

and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

```

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GHLJ
Center clone name: CH230-118C1
----- Summary Statistics
Assembly program: Atlas 3.0;
Consensus quality: 235051 bases at least Q40
Consensus quality: 237228 bases at least Q30
Consensus quality: 238719 bases at least Q20
Estimated insert size: 244886; sum-of-contigs estimation
Quality coverage: 8x in Q20 bases; sum-of-contigs estimation

```

FEATURES

```

/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-118C1"
1712..259601
/note="clone_boundary
clone_end:Sp6
site:EcORI
end_sequence: BH285254"
18349..119761
/note="wgs_contig"
complement(155798..262015)
/note="clone_boundary
clone_end:T7
site:EcORI
end_sequence: BH285252"
complement(261235..262017)
/note="clone_boundary
clone_end:T7
site:EcORI
end_sequence: BH285252"

```

ORIGIN					
Query Match	28.9%	Score 1655.4	DB 2	Length 263901	
Best Local Similarity	73.0%	Pred. No. 0			
Matches 2910	0	Mismatches 451	Indels 624	Gaps 36	

REFERENCE AUTHORS TITLE JOURNAL COMMENT	3 (bases 1 to 263901) Rat Genome Sequencing Consortium. Direct Submission Submitted (10-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA On May 10, 2003 this reference version replaced gi:24954086. The sequence in this assembly is a combination of BAC-based reads	Best Local Similarity 73.0%; Pred. No. 0; Matches 2910; Conservative	0; Mismatches 451; Indels 624; Gaps	36;
QY	1833 AGGGAAGAACCCAGGCTACCTTTAATATCATGAAACAGCGAG- -AGGGAAGCT	1891		
Db	263901 AGGGGAATACCCCAACTACTCTTTAATGAGAACTCATGAAATCAGCGGAAAGCGAAGCT	2638		
QY	1892 GGAGAGATGAGAAAGACCCCGGGGCGAACATGTAGAACGACAGCGAGTTGACGCG	1951		

Db 263841 GGAGAGGTAGAGGAATA-----AGACAAGGACAAAGCCAGTTGAGTGC 263798
Qy 1952 TCCGTGAATCAGCTCTGTGAGGACAGAGGAGGAGGCT 2011
Db 263797 TCTGTGGAATT-----CCGATATGAGCGCGGAACAGCAGAGGCT 263758
Qy 2012 AGGGTTAATGTGAGACAGAGGAAACAGAGGTAGACACAGGAAACAGACAGAGAGCGGGAG 2071
Db 263757 AGGGCTGATTTAAGACAGAGGAACTGAAGGTGGSCACAGGAAACAGACAGAGAGCGGGGAT 263698
Qy 2072 CCAAGTAAACAAGAAATGTCCTTCTCACTGTGGCCAGAGCGTCCATCT-GTGTCCACA 2130
Db 263697 CCAAGTAAACAAGAAATGTCCTTCTCACTGTGGCCAGAGCGTCCATCTGTGGGTCACACT 263638
Qy 2131 TACTCTGAATGTTTATCAGACCTGCAAGGCTGCTTGGGAGGAGCGGAAAGATATGT 2190
Db 263637 TACTCTGAATGTTTATCAGACCTGCAAGGCTGCTTGGGAGGCGGCTGAAGAGATATGT 263578
Qy 2191 GAGAGCCAGGGGAGACAAAGGGGCGCTAGAAAGGAAAGAGGAGCAACAGGCCACACA 2250
Db 263577 GAGAG-CGGGGAGACCTAAGGGGACCGAGAAAGGAAAGAGAGGCAACAGGCCACACA 263519
Qy 2251 AGAGGGCAGAGGCCCAAGAACTGAGTTAACTCTTCTTGTTCATCTTCCATAGAGGCAAG 2310
Db 263518 AGAGGGCAGAGGCCCAAGAACTGAGTTAACTCTTCTTGTTCATCTTCCATAGAGGCAAG 263462
Qy 2311 TGGGAACCTGTGACCAACATCCCCCAAGGCCCACTAACCCATACCAAGTTGGGCTG 2370
Db 263461 TGGGAACCTGTGACCAACAT-----CCCCCACTAACCTAACCAAGTTGGGCTG 263412
Qy 2371 AGTGGCACTTCTAGAGTCCCTGAGAGACAGAGCTGAGCTTGTCTCTTGGACCTGACCAA 2430
Db 263411 AGTGGCACTTCTAGAGTCCCTGAGAGACAGAGCTGAGCTTGTCTCTTGGACCTGACCAA 263352
Qy 2431 GCTGACCAATGTTTCTCAGTACCTTTGAAATGCTTCAAGAGCTTGAAGACAGGCACTGA 2490
Db 263351 GCTGACCAATGTTTCTCAGTACCTTTATCATGCCCCCTCAAGAGCTTGAAGACAGGCACTGA 263292
Qy 2491 CATTTTAGGCCATGGGCTAACCCCTGAGCTTGCACACAGAGGCTCAAGTGAACCTCCAGG 2550
Db 263291 CATTTTAGGCCATGGGCTAACCCCTTGCACACAGAGGCTCAAGTGAACCTCCAGG 263232
Qy 2551 GACACAGCTGACAGACAGGTGACCTTTATCCCAAGAGCAACCACTTGGCATAGTGGCT 2610
Db 263231 GACACAGCTGACAGACAGGTGACCTTTATCCCTGAGAGCAACCACTTGGCATAGTGGCT 263172
Qy 2611 GCAAATGGGAATGCAAGTTGAATCAGGTCCCTTCAAGAAATATGTCATGCAAGACTTAAAG 2670
Db 263171 GCAAATGGGAATGCAAGTTGAATCAGGTCCCTTCAAGAAATATGTCATGCAAGACTTAAAG 263112
Qy 2671 ACCCTGAGAGAGAGGGGTATGCTCTGCCCCCACCACCAATAGGGGAGTAACTATCTCT 2730
Db 263111 ACCCTGAGAGAGACTGTATGCTCTGACCCACCAACCAATAGGGAATGAAATATCTCT 263052
Qy 2731 AGGGGGCTGCGACCTTGGGAGACACCAACATTACTGAGTGTGAGCCCAAGAAAACT 2790
Db 263051 AGGGGGCTGCGACCTTGGGAGACACCAACATTACTGAGTGTGAGCCCAAGAAAACT 262992
Qy 2791 GACCGCCTGTGTCTCTGCCACCTTCACACTCTAGAGCTATATTAAGAGAGTACAGTA 2848
Db 262991 GACCGCTGTGTCTCTGCCACCTTCACACTCTAGAGCTATATTAAGAGAGTACAGTA 262932
Qy 2849 GATAGGGTGGAGCTGTGAGAGGGAGAGTGTCTGAGGTTGAGGGGTGAGGGGAAAGC 2908
Db 262931 GAT-AGGGGAGAGCTGTGAGAGGGAGAGGAGCTTCTGAGGTGAGGGGTGAG-AGGAAAGT 262874
Qy 2909 CAGAGCAGGGAGTCTGACTTGTCTCTTGAACAATGTCTACTTATGTTAAACAGCA 2968
Db 262873 CAGAGCAGGGAGTCTGACTTGTCTCTTGAACAAGTGTCTACTTATGTTAAACAGCA 262814
Qy 2969 TGAACCTGTTAAAGACCAACATCTAGACCTCTGAAAAAGACAGACCCCTGAGGACAG 3028

Db 262813 TGAACCGCTAAAGACCAACACCTACACCTCTGAAAAAACAGAGTCTCTGAGAGAGCG 262754
Qy 3029 GGTGTCTCTGAGCCCTTGGGTGCTTATGATGTCACAAAGAGAGGAGATGATGATAT 3088
Db 262753 GGTGTCTCTGAGCCCTTGGGTGCTTATGATGTCACAAAGAGAGGAGATGATGATAT 262694
Qy 3089 AAGGCCCAAGAGCGTTAGAGAAAGGCACTTGGAAAGGGTCAGTCTGCAAGAGCCCTAT 3148
Db 262693 AAGGCCCAAGAGCGTTAGAGAAAGGCACTTGGAAAGGGTCAGTCTGCAAGAGCCCTAT 262634
Qy 3149 CCATGGAATCTGAGAGCTTGGGCGCAACCTGTGTAATCTCTGGGCTTGCAGAGCAATTCAA 3208
Db 262633 CCATGGAATCTGAGAGCGGGGCAACCTGTGTAATCTCTGGGCTTGCAGAGCAATTCCG 262574
Qy 3209 AGCAGACCTGACATCTCTGAGAGCTTGGGAGAGCGGAAAGAGCAACCCCACTTAATA 3268
Db 262573 AGTAGACCTGACATCTCTGAGAGCTTGGGAGAGCGGAAAGAGCAACCCCACTTAACA 262514
Qy 3269 CCCTTCTCCCTGAGCCCAAGATTAAACCTTGTGCTTCCCTTCCACCTCCATC 3328
Db 262513 CCCTTCTCCATCAGACCAAGATTAAACCTTGTGCTTCCCTTCCACCTCCATC 262454
Qy 3329 AGGAGTGAAGGTTGAGAGGAGGAGGTTAAACCTTCAATGTCCAAATCATGTGACAG 3388
Db 262453 AGGAGTGAAGGTTGAGAGGAGGAGGTTAAACCTTCAATGTGACAGTGAATGTTAAAGTCT 262402
Qy 3389 ATATATGATCAGATATGATGATGAGGACAGAAAGAAATCTGCA-----GACTTAATG 3442
Db 262401 ATATATGATCAGATATGATGATGAGGACAGAAAGAAATCTGAGTTAAATGTTAAAGTCT 262342
Qy 3443 GTTAAATGTGTA--GTCTGTGCAATGTGTGTCTGACTGAAACCGGCAATGCT 3499
Db 262341 GTCTGTGTGTGATGATGATGATGATGATGATGTGTGTGTGATGAAAGGGCAATGAT 262282
Qy 3500 GTGACCTGTGCACTCTGTGCGGTGAGGTTAACAGCTGAGGTTGTGTAAATTTGCC 3559
Db 262281 GTGACCTGTGCACTCTGTGTGATGATGATGATGATGATGATGATGATGATGATGAT 262230
Qy 3560 CAAGGCAAGTGGGTGAATCCCTTCCATGATGTTAAAGATGATGATGATGATGATGAT 3619
Db 262229 -----AAGTCATGAGAACCTCCATGATGTTAAAGAAATGATGATGATGATGAT 262177
Qy 3620 CAAGACCATGAGAAATAGATGACACTGTATATGTCTCTTAAGCTAAAGTACAGAG 3679
Db 262176 CAAGGCAATGAGAAATAGATGACACTGTATATATGTCTCTTAAGCTAAAGTACAGAG 262117
Qy 3680 TCTTTGAGAGACCTGTCTAGAGATGTGGGCAACAGACTACAGACGATCTGAC 3739
Db 262116 TCTTTGAGAGACCTGTCTAGAGATGTGGGCAACAGACTACAGACGATCTGAC 262057
Qy 3740 GAGTAAAGAGAGAGAGAGGGGTGTGAATTTCTTTACTATCAAAAGGAACTGAGTGC 3799
Db 262056 GAGTAA--AAGAGAGAGTGAAGGTATAGAAATTTCTTTACTATCAAAAGGAACTGAGTGC 261999
Qy 3800 TGCACCTGCAAGATGATGCTTCTCTAGACATCATGACTTTGTCTCTGGGAGCCAGCA 3859
Db 261998 TGCACCTGCAAGATGATGCTTCTCTAGACATCATGACTTTGTCTCTGGGAGCCAGCA 261939
Qy 3860 CTGTGGAATCTGAGTCTGAGAGTGAAGAGGCTCCCTCAGGCTGAAGCTATGACAGATA 3919
Db 261938 CTGTGGAATCTGAGTCTGAGAGTGAAGAGGCTCCCTCAGGCTGAAGCTATGACAGATA 261881
Qy 3920 GCCAGGTTGAAA--GGGGAGAGGAGAGCTGGAATGGAGGCTTGTGTGTTGAAGCAG 3977
Db 261880 GCTTAGGCTGTAAGAGAGGAGAGGAGGAGCTGGAATGGAGGCTTGTGTGTTGAAGCAG 261821
Qy 3978 GGAACAGATATTAAGCTGGAAGAGAGGAGCTTAACTCAAGT-TGTTCAACTCACCT 4036
Db 261820 GGAACAAATATTAAGGCGCTTAAGAGAGGAGGAGCTTAACTCAAGTGTCTCAACTCACCT 261761
Qy 4037 TCAGATTAAATAATACAGAGTAAAGGCC---TGGGTGAAGGAGGTGTGAGAGCGCTC 4093
Db 261760 TCAGATTAAATAATACAGAGTAAAGGCCATGTGGGTGAAGGAGGTGTGAGAGCGCTC 261701

```

QY 4094 CTGCTCTCCTGTCAGTGGCCCTTGAGCCCTTTGGGGAGGAGAAATGCGCCAAAGACTAA 4153
Db 261700 CTGCTCTCCTTATATCTGCGCAATCGGCGCTTTGGGGAGGAGAAATGCGCCAAAGACTAA 261641
QY 4154 AAAAAGCCATGAGCCAGAGGGGCGAGGCGAACAGACTTTATGGGCGAAACCTTTGGG 4213
Db 261640 AAAAAGCCCTGAGCGAGGGGCGAGGCGAACAGACTTTATGGGCGAAACCTTCAAGG 261581
QY 4214 CCCGTAGTATGATGACAGAAATCTGCCAATGATACCTTCTTCTTCTTAACGAGACG 4273
Db 261580 C-----TGCTGTCTCTCTGTCACCTTCAAGAGCC 261553
QY 4274 GAGGSACTGAGTTTACCACTCCCTATCAGTATAGAGAAAGTGAAGTGAAGTTTAC 4333
Db 261552 AAGGATCAAAAGAGAGGAGCGAGCAGAGAGGATGGAGGAG----- 261507
QY 4334 CACTCCCTATCAGTATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4393
Db 261506 ----- 261507
QY 4394 GAGAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4453
Db 261506 ----- 261507
QY 4454 GTTTACACTCCCTATCAGTATAGAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4513
Db 261506 ----- 261507
QY 4514 GTGATAGAGAAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 4573
Db 261506 ----- 261507
QY 4574 AGTCAGCTCGTATCCAGCAGAGAGCTCCAAATTATAGCGACGAGCATATGGATGGAT 4633
Db 261506 -----GGTCCAGCAGATGACTCCAAATTATAGCGACGAGCATATGGATGGAT 261457
QY 4634 ATAAAGGGGCTGAGCACTGAGAGCTGTGAG-----AGATTCTCCAAACCCAGGTAAGAG 4689
Db 261456 ATAAAGGGGCTGAGCGCTAGAGCTGTGAGAGCCAGATTTCTCCATCCAGGTAAGAG 261397
QY 4690 GAGTTTCGAGTGGGGGCTCTTCAACCCACAGACCTCTCCCACTAGAGAAAGTGC 4749
Db 261396 GAGTTTACGCTGGGGGCTCTTCAACCCACAGACCTCTCCCACTAGAGAAAGTGC 261337
QY 4750 CTTTCTGAGAGTGGGGTTCAGGCGGCTCAGAGATTCAGAGGGTGCCTTCAACAGCC 4809
Db 261336 CTTCCCTGAGAGTGGGGCT-----CTTCCACAGGCG 261305
QY 4810 TGGGAAGTTCTCAGTGGCAGAGAGTTTCCACAAAGAACTATGATGCCCCCTTACG 4869
Db 261304 TGGGAAGTTCTCAGTGGCAGAGAGTTTCTA-----CTGATGCCCCCTTACG 261254
QY 4870 CTGCTCTCCATCTCTCTCTGAGAGTCTCTCCCGCTTGGTTATCTTGGCTCT 4929
Db 261253 CTGCTCTCTCTCTCTCTCTGAGAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 261214
QY 4930 CGTCTTCAAGAGATTGCGCTGTCTGCTCACTCACTCTCTCTCTCTCTCTCTCTCTCT 4989
Db 261213 AGGTTCAGAAAGATTGCGCGGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 261160
QY 4990 TTGCTTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5049
Db 261159 -TGCCTTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 261101
QY 5050 CTCTCTATTTGATTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5109
Db 261100 CCCCCCTTATTTTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 261070
QY 5110 TTCTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5169
Db 261069 -----CTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 261061

```

```

QY 5170 AGTGTGAGAAATCACACCTGGGGTTCCAGCC--TTATGTAAACATCTTCCAGTAGGCC 5227
Db 261060 AATGCTGGGAATCAAAACCCAGGGCTTTATACAGTCAAGTAAAGCATCTCCAGTAGGCC 261001
QY 5228 AAGCTTCACTGCTGTGGGTCTCTTAACTTCTCTCAACCCCTGGCTTTGCTTCTTC 5287
Db 261000 AAGCTTTATCTCT--CTGGGTCTGTCTTAACTTCTCTCACTCTCTGTCTTGTCTCTTC 260942
QY 5288 ATCTGTGTAGATCTCTAGATTTGGTCTCCAGCCCTGTCTCTCTCTCTCTCTCTCTCT 5347
Db 260941 GTCTTAGTGAAGATCTCT--GGTCCGTCTCTCAAGTCTGTCTCTCTCTCTCTCTCTCT 260883
QY 5348 CCTCTCTGTCTCACTGCGCAGCTGTGGTGTCTGTCTCACTGTGTCTCACTTCTTC 5407
Db 260882 CTCTCTCTGTCTCACTGCGCAGCTGTGGTGTCTGTCTCACTGTGTCTCACTTCTTC 260823
QY 5408 AGGATTTCTGTAAAGTTAAACAGGTGAATGTTTCCCTGTGAACAGAGATCAAGAT 5467
Db 260822 GTGAAAAGATTAAACGGGTGAAGATGCCCCAGTTTCCCTGTGAACAGAGATCAAGAT 260763
QY 5468 TCT-CCCCGAATTCAGGCTTTCAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5526
Db 260762 TTTCCCGAAGCCAGACTTTCAGACCCGCG--CCTGTCCAGACACTGAGACTTTA 260705
QY 5527 GCAAACCTCAGGCACTTACCCCACTAGACC----- 5559
Db 260704 GCAAACCTCAGGCACTTACCCCACTAGACC----- 5559
QY 5560 -----TGTGACAGAGACAGGCACTTTTAA 5585
Db 260644 GACTTCGAGTTTCTGCAAGAGAGAGAACTTTGTGTGTGGGAGAGGCACTTTTAA 260585
QY 5586 TGGAGTCTGTGGAGAGAGCCATAGGCTGCTGTGTAAAGAGGCGGAGAGTGTGTCT 5645
Db 260584 CGAGTCTGTGAGGAGAGCCATAGGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 260525
QY 5646 AGAAAGTCAAGACTTCAATAGAGCTTGAAGCCACACCAAGAAATGACAGAGATCCCT 5705
Db 260524 -----TGAATTTCAACAGAGAG--CTAGCAGACACAGAGAT-----GACATATCCCT 260479
QY 5706 CCTATCTCCCCCATTAAGAGTTTGAAG 5730
Db 260478 CCTATCTCCCCCATTAAGAGTTTGAAG 260454

RESULT 8
HAMSHCA
LOCUS
DEFINITION HAMSHCA 32415 bp DNA linear ROD 17-FEB-1995
1-39, complete cds.
ACCESSION L15351
VERSION L15351.1 GI:402373
KEYWORDS alpha-cardiac myosin heavy chain.
SOURCE Mesocricetus auratus (golden hamster)
ORGANISM Mesocricetus auratus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
Mesocricetus.
REFERENCE
1 (bases 1 to 32415)
AUTHORS Wang,R., Sole,M.J., Cukerman,E. and Liew,C.C.
TITLE Characterization and nucleotide sequence of the cardiac
JOURNAL alpha-myosin heavy chain gene from Syrian hamster
MEDLINE 5. Mol. Cell. Cardiol. 26 (9), 1155-1165 (1994)
PUBMED 95115033
COMMENT Original source text: Mesocricetus auratus (strain F1B) male adult
liver DNA.
FEATURES
source
1..32415
/organism="Mesocricetus auratus"
/mol_type="genomic DNA"
/strain="F1B"
/db_xref="taxon:10036"

```

```
misc_difference 1. 60
/sex="male"
/cisue_type="liver"
/dev stage="adult"
/strand_name="overlapping region"
/notice=""_in misc_difference; The first 60 bp of this
sequence overlap with the last 60 bp of the sequence under
Genbank accession number J12104. Their combined sequence
spans a 66,318 bp genomic region."
/replace=""
4000. .4005
4032. .4037
join(4064. .4083,4519. .4689,5146. .5363)
/strand_name="5' untranslated region"
/notice=""The 5' UTR for hamster cardiac alpha-myosin heavy
chain gene contains the first two exons and the first 17
bp of the third exon; putative"
/function="5' UTR for hamster cardiac alpha-myosin heavy
chain gene"
exon 4064. .4083
/number=1
4084. .4518
/number=1
4519. .4689
/number=2
4690. .5145
/number=2
5146. .5363
/number=3
join(5163. .5363,6526. .6669,6893. .7049,7168. .7195,
7412. .7523,7761. .7853,8540. .8603,8870. .8968,9337. .9440,
9531. .9669,10084. .10352,11049. .11219,12539. .12848,
13363. .13433,13889. .13976,14060. .14177,14299. .14422,
14589. .14725,17338. .17593,17840. .18082,18255. .18431,
18950. .19095,19295. .19385,21101. .21490,21820. .21946,
22037. .22155,22444. .22640,22183. .23366,23503. .23668,
23762. .23866,24640. .24948,25342. .25545,27609. .27734,
27843. .28118,29407. .29502,29617. .29751,30113. .30136)
/notice=""putative"
/codon_start=1
/product="alpha-cardiac myosin heavy chain gene"
/protein_id="AAB59701.1"
/db_xref="GI:402374"
/translation="MTDSQMADEFGAAAEYLRKSEKRLLEAOTRPPIRTETCFVPPDKE
EFVYAKIVSREGKVTAETENGKTVLEQVQWQNPQPKDIEMAMLTFLBPAVL
YNLKERAYAMNIVYYSGLFCVTNPKYKMLPYNAEVAARAKRSSEAPAHFISIN
AYQWMLTDRNOSTLITGESGAGKTVTKRVIQVPAIAIGRSKKDNNAKGTLE
DOIQANPALEAFGNATVVRDSSRGKPIRIHFGATGCLASADITETLLSEKRVF
QLKERNYHIFQILSNKPELDMVLVTNNPYDAFVSQGEVSAVIDSEELATD
SAPDLGLPTAEKAGVYKLTGALMAYGNMKFKQREQAEPPDGETADDSAYLMGIN
SADVLKGLCHPRVAVGNEVYTKQSVOQVYYSIGALSKSYEKMFMMVTRINATLET
KOPROYFIGVDIAGFEIFDPSPEOLCINTFNKLOQFPNHHFVLEOEYKKEGIE
WEPTDPMDOIACIDLI EKPMGIMSLIEECMPKATDMPFKALYNHIGKNNPKX
PRYVKQGEAHFSLVHAGTYDNIILGMLBKNDPLAETVGLYQKSLKMTLTET
YASADGDSGKGGKGGKGSFQTVSALHRENLMKLTNTLTPHFRCLIPERKA
PGVNDNPLVMHQLRCNGVLBGIRICRQGFPRILLYGDPORRYRLNPAALPEQFIDS
RKGAELKLSLSDIDHNOYKFGHTVFFKAGLIGLEMBERLSRIITRIQAOARGOL
MRIEFKWVERRDALVIQNNIRAFMGVKNPMWLYFKILPKLSAETEKEMANME
BFGVSESLSEKSEARKELEKMYKLSLOEKNDLOFOAOADONINDAEECDOLIKK
IQLEAKYKMTERTLEDEBEMAEITSKRKLDEDCSLKCIDDLLETITAKYERKA
TENKVKULTEMAGLDETIAKLTKEKALOAHQOALDDLOAEEDKVNITLTKSVKKA
QOVDLGSLEQEKVRMDLERARKLEGLDNLVQESIMDLENDKLOLEBKLEKEDD
IQOONSKIDEQALALQLOKLEKNQARIELELELEAERTARAKVEKLSIDLTRIE
EISERLEBAGATVOIEMNKKRAEPQKRRDLEBTLQHEATAALRKHADSVSE
LGEQIDNLQRYKQKLEKESRFXLELDQVSNMOQIIKAKANLEKVSRTLEDQNEVR
VKLESQBSLNDFTTORAKLOTENGELARQLEEKALISQITRKSLSYTOAMELKQ
LEBEGKANALAHALQSRHDCDILRQYEEEMAKALQKVLISKANSVAYQMTKE
TDAIQRTEELEBAKKLAQRLODAEBAVEAVANAKSLEKTRKLONEIDLDWYER
SNAALALDKQRPNDILAEWKQYESQESQESQESQESQESQESQESQESQESQESQ
EHLETFKRENKQLEBEISDLBOGEGKNVHELEKRYKOLEVEKMELOSLEBAESL
LEHEGKILRAOLFQNOIKAEIERKLEKDEEMOAKRNHLRVVDSIQSLDATTERR
NEARVYKKMGDINMEWETIQLSQNRILASBAQKILKNAQAHKDTQQLDDALHAND
LKENIATVERNTLLQAELEBLRAVVEQTERSEKLAQBELITSERVOLHSONTSIL
```

```
introns
exon 5364. .6525
/number=3
6526. .6669
/number=4
6670. .6892
/number=4
6893. .7049
/number=5
7050. .7167
/number=5
7168. .7195
/number=6
7196. .7411
/number=6
7412. .7523
/number=7
7524. .7760
/number=7
7761. .7853
/number=8
7854. .8539
/number=8
8540. .8603
/number=9
8604. .8869
/number=9
8870. .8968
/number=10
8969. .9336
/number=10
9337. .9440
/number=11
9441. .9530
/number=11
9531. .9669
/number=12
9670. .10083
/number=12
10084. .10352
/number=13
10353. .11048
/number=13
11049. .11219
/number=14
11220. .12538
/number=14
12539. .12848
/number=15
12849. .13362
/number=15
13363. .13433
/number=16
13434. .13888
/number=16
13889. .13976
/number=17
13977. .14059
/number=17
14060. .14177
/number=18
14178. .14298
/number=18
14299. .14422
/number=19
14423. .14588
/number=19
14589. .14725
/number=20
14726. .17337
```

exon /number=20
1738. .17593
/number=21
17594. .17839
Intron /number=21
17840. .18082
/number=22
18083. .18254
Intron /number=22
18255. .18431
/number=23
18432. .18949
Intron /number=23
18950. .19095
/number=24
19096. .19294
Intron /number=24
19295. .19385
/number=25
19386. .21100
Intron /number=25
21101. .21490
/number=26
21491. .21819
Intron /number=26
21820. .21946
/number=27
21947. .22036
Intron

Query Match 25.2%; Score 1444.8; DB 10; Length 32415;
Best Local Similarity 69.3%; Pred. No. 0;

Matches 2854; Conservative 0; Mismatches 932; Indels 330; Gaps 52;

QY 226 AGAATAAGCAATTTTCTTAAAGCCAAATCTGCTCTAGACTTTCTTCTCTGACT 285
DB 1 AGAATAAGCAATTTTCTTAAAGCCAAATCTGCTCTAGACTTTCTTCTCTGACT 60
QY 286 CGGTCCCTGGGCTCTAGGGTGGGAGGTGGGCTTGAAGAAGGTGGGGAAGTGA 345
DB 61 CCATCTGGGGCTTTGAGTGGGAGGTGGGCTTGAAGAAGGTGGTCTTCAAGAG 120
QY 346 AAGCCGATCCCTAGGAGCCCTGTGAAGTTGAGAGCTTCCCTGTACAGCATGGCTCATA 405
DB 121 CAAAGGCTCCCGAGTTCCTGTGAGCTGTGAGCTTCCCTGTGACAGCATTGTGTGAG 180
QY 406 GATCTCTTCCAGCCAAACATAGCAAGAGTATCTCTTTGTGACTTCCCGAGGCC 465
DB 181 GACCCCTCCCA-CAGAAATACAGCAAGAGTGGCACACGCT-----CCTTCCCGGCC 234
QY 466 AGTACTGTCAAGTTAAACAGATTTAGAGAGCTTGAAGCTTGAAGCTGAAGCTGAAG 525
DB 235 AGGGCTGTGAGATGAAA-TGGGTTCAAGGGGCTGTGAATCTCATGTGGCTGTAG 293
QY 526 CTCATCCACCAAGCAAGCACTAGTGCCTGCTAGTTAGTATCTTACGCTGATATAT 585
DB 294 C-----CTCTGTGACATTACTTACACCGGTTCTATCCGTACTCAGACTGGCAGAAC 349
QY 586 GCAGAGCTGGGCAAGAG--TCTGGGGTGTAGAACTGACAGTGAATTTTCAGTCG 643
DB 350 CCAGAGCTGGGCAAGAGAGGAGGCGGGAATTAAGGCTGACCAATGACTTTTCAGTCT 409
QY 644 GCAAGGTATGACCCCTCAGAGATGTAGTATGTCCCTTGA--TCCATCCAGGC 701
DB 410 GCGA--GTGTGACCCCTCGGAGATGTGTAAAGCTCTCTGAGACCCCTCATCTTGGT 467
QY 702 AGGTCTTAAAGAGACATGGAGTGAAGATGTAGTGTGATTCCTTCAACACAGCTAT 761
DB 468 AGGTCTTGAAGAGACACAGATCAAGAT--GTCAATGCCATTCCTCAACACAGCTAT 526
QY 762 CCAGAGTGTCTTGGCCCTTCACTTGAAGAGAGAGAGTAACTTAACTTATCTTTC 821
DB 527 CCAGTGTGGCCATGTCTC-TCACATATATTAGAGAG-CAGTAACTTGGCTATCTTTC 584

QY 822 TTCCTCCCATCTCCAGACACACCCCTGTGTGAGTATTCATTT-----CTTCC 875
DB 585 TTCCTTCTCCCATCCCTCTGTGTCCCATCCAGACCAACCCCTTGGCTGTATCTCAG 644
QY 876 TTCACGTCCCTCTGTGACTTTCATTTTGAAGGCTTTTGAACCTTGCAGCTCTGAGAGA 935
DB 645 TTCCTCTCCCATTTTCCCTCCATCTGCAAGGCTTTTGAACCTTGCAGCACTGAGAGA 704
QY 936 TAGAGTTTGGCCCTTGTGTGCAAGCAAGCTCAAGAGAAACACACACAGGGGAGCA 995
DB 705 TGGGATTTGGCTTTATGTGGGAAAGCCATTCAGAGAGACAGCCACAGAGAACCA 764
QY 996 GATTTTGAAGATCAGGAATCAATCACTGGCGGGCTTGGGGGTAGAAAAAGAGTAG 1055
DB 765 GATTTTGAAGAGAGAGACTTAATCATTTAGAGGGCCATAGGTAGAAAGAGAACAG 824
QY 1056 TGAATCCGCTCAGCTTAAGCCAAAGTAGTCCCGAGATATCTGCAACAGCTGGGCTGT 1115
DB 825 CAAGTCTCTGACAGCTTAAGCAAGCTAGTCCCTGGCAATCTGCTCA---GGGCTGCT 881
QY 1116 CCGGTAAGCTTAAAGAAATGTGGGTGAAAGCAATGGGATTGGAACATCTCT--- 1171
DB 882 AAGGCAAGGTTAGAAATGTGGGTGAGAGCAATGGGATTTTAATCATCTCTGGA 941
QY 1172 -----TGAGTCTCCCTCAACCCCACTTACAGACACACTCGTGTGT 1212
DB 942 ACCTGTTCAGCCCTGGGCGCAATCTCCCTTCAACATATCTTATGACATTTGTGCA 1001
QY 1213 GGCCAGACTCTGTTTCAACAGCCCTGTGTCTGACCACTGAGCTTGAAGCAACAGAGA 1272
DB 1002 GGCCGAGACTCTGTTTCAACACTCTCCCTGGGCTGTGATCACTGAGAGCAACAGAGA 1061
QY 1273 TGGGCTCTGTGTGAGATGAAGATTTTCAATGCAAAACAGAGGGGGAGG 1332
DB 1062 TGGGCTCTGTACTGAGTAGAG--GGTGTCCAAACAGCAAAACAGAGGGGGAGG 1117
QY 1333 AACAGAGAACGAATTAAGGAAGAAAGAAAGGACCAATCATGATGATGAGAG 1392
DB 1118 AACAGAGAACCATTAAGGAGAA----- 1142
QY 1393 AGATGGGAAGCCACACACAGCTTGAAGAGAGAAACAGAAAGGAGATTTGGGCA 1452
DB 1143 --AAGGAATGTCTTAACTCAGATTGAGCTGGAGAGCGGAGATGCTGGCA----- 1192
QY 1453 TGAAGGCAAGAAAGAGAGCCAGCCCAAGTCTCTTTATACCTCATCC 1512
DB 1193 ---GAGGCAACAGAAAGAGAGCCCAAGCCCACTTCTTTCATCAATCCC 1248
QY 1513 CGTCTCCCAATTAAAGCCACTTCTCTTATGATGAGCTGAGCTGAGGAGAGAGCC 1572
DB 1249 TTCCTCTCTTAAACCATTCCTTCTTAAAGACAGCTTGAACGCAAGAGAGATC 1308
QY 1573 CGTAG-----GAGAGATCACCTGATGAAGAGATGTGTGAGAGTCCAGGCAAC 1626
DB 1309 AGTATGAGAGAGAGTACACACGAGAGAGATGTGTGAGAGGCCAGAGAAAGCT 1368
QY 1627 TTAAGGCCAGACCTTAAAGAGCAAGATTAAGGTCTTCAAAGTGGCCAG--CTGTG 1685
DB 1369 AAGAGCCAGAGCTTAAAGAGCAAGATGAAGTGCACAAAGTATGCCCAGAGCTGGG 1428
QY 1686 CACACAGAGAGGTGAGAGCTGTGTGAGAGCTCAAGATTAAGATGATGTGAGATGGG 1745
DB 1429 CTTGAACAGAGTCAAGAGAGCGGTGAGAGCCC-----AGCAGAGAGCGCTTAAGATGCA 1483
QY 1746 CCGGGGGGGGAGTTCTGGGGGGGAGAGAGAGGTGAAGAGAGCTTGAACAGAGAT 1805
DB 1484 CGAG-----GAGAGCAATTTGGGAAGGGGCTTGAACGAGAGAT 1521
QY 1806 CTGGAAGGCTTGAAGACATTAAGAGAGAGAGAGAGAGAGCTTCTTAAATGTAAAT 1865
DB 1522 CTGGAAGGCAAGGAGAGTGAAC--TGCAAGGGAGAAACCTCAGGCTACTTT--GATGTAAAT 1578
QY 1866 CATGAAGACAGAGAGAGAGAGAGTGAAGAGATGAAGAGAGAGAGAGAGAGAGAG 1925

Db 1579 CACGAA-----GCCAAGGAAAAAGAGGCTGAGAGAGAAACCCTGGGCAAGACATGG 1631
Qy 1926 AAGCAAGACAGACGACAGTGTGAGCGCTCCGTGAAATACAGCTGTGTGAAGCAGAGCCCTGG 1985
Db 1632 -----TAGCAAGGCGGACGACCTCTGTGAGATCTGTCTGAGAA--CAAGAGCCCTG 1680
Qy 1986 GTATGACACACAGAACAGACAGGCTAGGGTTAATGTGAGAACAGGAAACAGAGGTAGA 2045
Db 1681 GTGGAGCGGCAAGAGCAGCCAGATC-----TGATTGCCACGGAAACAGAAAGGTAGA 1731
Qy 2046 CACAGAAACAGACAGAGACG--GGGAGCGCAGGTAAACAAGGAAT----- 2088
Db 1732 CCCTGGAAACAGCAAAACGACGAGAACCGAGGAAACAAAGAGCAGCTGTGCTCAAG 1791
Qy 2089 -----GGATCCCTTCACCTGTGGCAGAGGGTCCATCTGTGTCCACATAC 2133
Db 1792 CCAAGCTTGCCCTTGGGTGACTTTCCTGTGACCAAGGCTCCATGTGCTTCCACATAC 1851
Qy 2134 TCTAGAAATGTTCAATCAGACTGCAAGGCTGGCTTGGAGGCAAGCTGAAAGATGTGAG 2193
Db 1852 TCTAGAACTTTCATCAGACTGCGGGGCTG-----GAGAGCTCTGAAGATGTGAG 1903
Qy 2194 AGCCAGGAGACAGAGGAGGCTTAGAAAGAGAGAGAGGCAAAACAGGCCACACAGA 2253
Db 1904 AGCCAGAGAGAGAGAGGAGGCTAGAAAGAAAGAGAGGCAAAACAGGCCACACACA--- 1960
Qy 2254 GGGAGAGGCCCAAGACTGAGTTAATCTCTCTGTGTGCAATCTTCCATAGAGAGCAATGG 2313
Db 1961 ----AAAGCCAGAGCTGAGTTAACTCTTCTCTGT-----CTCCACAGAGACAAATGG 2010
Qy 2314 GAATCTGTGACCAACATCCCACTGAGCCCACTACCAATCAAGTTTGGCTAGT 2373
Db 2011 AGACTCGGTGCCCAATCTCCCAATCCCACTACCAATCTTGTCTGTCT 2070
Qy 2374 GGCATCTAGGTTCCCTGAGACAGAGCGCTTGTGTCTTGTGAAGCTGAGCCAAAGT 2433
Db 2071 GGCACAGGAGGTTCCCTGAGGACAGAGATAGTCTTGTGCACTTGGACCTGACCAAGGCT 2130
Qy 2434 GACCCATGTTCTCAGTACCTTTGAATGCCCTCAAGCTTGAGAACACAGGCAAGTGCAT 2493
Db 2131 GACCCAAATGTTCTAGTCCCTTATCATGTCTCTGAGACCTTGAACACAGGCAAGTGCAT 2190
Qy 2494 ATTAGCCATAGGCTAAACCTGAGAGCTTGACACAGAGGCTCAAGTGAAGCTCCAGGAG 2553
Db 2191 ATTAGGCCACAGGCTAACCTGTGACTTGGACAAAGTGAAGCT-----TCCAGGAGC 2241
Qy 2554 ACAGCTGACAGACAGGTGCTTTATCCCAAGAGCAACATTTGGCAATAGGTGGCTGCA 2613
Db 2242 CTAGCTCAGACAGAGTGGCTTGCATCCCTGAGAAACAATCATTTGGCATAGTCACTGCA 2301
Qy 2614 AATGGGAATGCAAGGTGGAATCAGGTCCCTTCAAGA-----ATACTGATGCAAAACCTAA 2669
Db 2302 GATGGGAATACAAAGTGTGCTCAAGTCCCTTCAAGAAACTAATCTGATATATAGCTAA 2361
Qy 2670 G-----ACCCCTGAGAGAGAGGGGTATGCTCTGCCCACTCCACCACTTAAGGGAGT 2720
Db 2362 GTGTGTGTGGGTGCTGAGAGAGTGTGTATGCCCTGAGCCCAATTTGGGCAAGAGGT 2421
Qy 2721 GAATATCTTAAGGGGCTGCGACCTTGGAGAGACCAACATTTAGAGTGTGAGCC 2780
Db 2422 GTCTTA-----GGGGCTGTGAGACTTGGGGGACACGCG--TCACTGAGATTTGCTGATCC 2474
Qy 2781 CAGAAACATGACCGGCTGTGTCTGTGCCCACTCCACACTCTAGAGATATATTGAAGG 2840
Db 2475 CAGGAATTTGACCGCTGTGTCTTCACTCCACCTCAACTCTAGAGATATATTGAAGG 2534
Qy 2841 TGACAGTATAGAGGTGGAGCTGTGTAGCAGAGAGAGTGTCTCGGGGTGAGGGTGTAG 2900
Db 2535 TGGCAGTATGATAGGTGG-----GTAGTTCTCGGGGTATGACAGTGTAG 2578
Qy 2901 GGGAAAGCAGAGCAGGGAGTCTGGCTTTGTCTCTGAAACAATGTCTAATTAGTTAT 2960

Db 2579 GGG--AAGCAAGAAAGAGGAGTCTGACTTTGTCCCTGAAACAATGTCTCTTAATTAT 2637
Qy 2961 AACAGGATGACCTGTCTAAAGACCAACATCTACAGCTCTGAAAGAGACAGACCCTGG 3020
Db 2638 AACAGGATGACCTGTCTAAAGACCTTAACCTACAGGCTCTGAGA--AATGCTGTCTGA 2695
Qy 3021 AAGCAGGGGTGTCTCTGAGCCCTTGGGTGCTTGAATGTGTCCACAAAGAGGGCATAGT 3080
Db 2696 AAGCAGAGGAGAACTCTAGTGTCCGGGTGCTTGAAGGTGCCACAGAAATGTTGTATTA 2755
Qy 3081 GTAGATTAAGGCCCCAGAGGCTTA--GAGAAAGGCACTTGGGAAGGGGCACTGTGCA 3139
Db 2756 AGCCCCAGAGAGGTTAGAGAGGGGTACAGAGTTTACCTGGGAAGGAGAGATGTGGCG 2815
Qy 3140 AGCCCTATTCATGGAATCTGAGCTGTGGGCCAACTGGTGTAAATCTGTG----- 3191
Db 2816 AGCCCGCAGCAGTGGAAATTCGAGGCTTGGGGCCAAAGGATTAATCTCTGAGCTGCCA 2875
Qy 3192 ----GCTGCCAGGCAATTCAAAGACACCTGCATCTTGGCAGCTTGGGAGGCTGAA 3247
Db 2876 GTCCCCAGCAGCCATCCCAAGCGGACCTGCATCTCTGGCAGCCTGGAAGAGACGAA 2935
Qy 3248 GGGAGCAACCCCACTTATACCTTTCTCCCTACAGCCCAAGATTAACCTCTGGCC-- 3306
Db 2936 GGGAGCAACCCCACTTACACCCCTCTCTCAGCCCTTAGGATTAACTCTGTGCT 2995
Qy 3307 TTCCTCCCTCCCACTCCCATCAGAGTGAAGGATTGACAGAGGAGGATTAACCTTACA 3366
Db 2996 TTCCACCTCTCATCCCCATACAGAGTGAAGGTTACAGAGGAGGTTAAACCTGTG 3055
Qy 3367 TGTCCAAACATCATGTGTGACGATATATGATCAGTATGTGTAGAGGCAAGAAAGAAAT 3426
Db 3056 TGTCCAAACATGTGTGTGCA--GATATATGACACAGTACATGTAGACGCAAAAAAGAAAT 3114
Qy 3427 CTGACGCTTAAATGGGTTATGTGTAAATGTC--TGTGTGATGTGTGTGTCTGACTGA 3485
Db 3115 CTGCAAGCTTAACTGAGGTAAATGTGTGAAGCGCTGTGTGTGTGTGTGTGTGTGT 3174
Qy 3486 AAACGGGCAATGTGTGACGCTGTTCAGTTCTGTGAGTGAAGTTACAGACTCAGAGTTT 3545
Db 3175 GTGTGTGTGTGTGTC--GCACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3233
Qy 3546 GTGTGTAAATTTGCCCAAGGCAAAAT--GGTGAATCCCTTCATGTTTAAAGATTTGA 3604
Db 3234 GTTTGTAAATTTGCCCAAGGCAAGTGGGGTGAATCCCTCCGTTGTAAAGACATTTGA 3293
Qy 3605 TGATGCTGTGATCTCAAGGACCATGAGAAATGAATGACATCTATATGTGTCTTA 3664
Db 3294 TGATGG--CGGCATGTGGGGACAAAT--CAGGTAAAGTAACTCTATATGTGTGTGT 3351
Qy 3665 GCTAAGGTAGCAAGTCTTTGAGAGCACTGTCTAGAGATGTGGCAACAGAGACTACA 3724
Db 3352 ACCAAGGTAGCAAGGCTTTGGAGGATACCTGTCTAGAGACATGGGCAACAGAGACATA 3411
Qy 3725 GACAGTATCTGTAACAGATTAAGGAG-----AGAGAGGGGGT 3763
Db 3412 GACAGATTTATACAGGTAAAGGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3471
Qy 3764 GTAGAAATCTCTTACTATCAAGAGAACTGAGTGTGTGACCTGTGCAAAAGTGTCTTC 3823
Db 3472 AGTAAATTTTATTAACAAAGAACTGAGTGTGTGAACTGTGAACTAAAGTGTCTTC 3531
Qy 3824 CTTAGACATGATGACTTTGTCTGTGGGAGCCAGCACTGTGAAACTTCAAGTCTGAGAGA 3883
Db 3532 CTTGAGACATTTTATCTTTGACTCTTGGGATCTAGCACTGTGAACTCTGAGAGG 3591
Qy 3884 GTAGAGGCTCCCTCAGGCTGAGCTATGACATAGCCAGAGGTTGAAAGGGGAAAGGA 3943
Db 3592 GTAGAGAGATCCCTCAGGCTGAGGCTGTGCAATAGCCAGAGGAGCGGGGGA----- 3646
Qy 3944 GAGCTTGGATGGAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4003
Db 3647 -----GGAGCTGT 3695

QY	5147	CCCTTCCTTCCTTCCTTCGTGTCAGAGTGCTAGGAATCAACCTGGGGGTTCCACCCCTATG	5206
Db	1089	TCCTTCCTTCCTTCCTTCGTGTCAGAGTGCTAGGAATCAACCTGGGGGTTCCACCCCTATG	1148
QY	5307	TAAACAATCTTTCAGTAGAGCCACAGCTTCAGTGTCTGGGTGTCTCTTAACCTTCTCA	5266
Db	1149	TAAACAATCTTTCAGTAGAGCCACAGCTTCAGTGTCTGGGTGTCTCTTAACCTTCTCA	1208
QY	5267	CCCCCTGGCTGTCTCTGTTCATCTCTGGTCAGGATCTTAATTTGTCTCCAGCTCTG	5326
Db	1309	CCCCCTGGCTGTCTCTGTTCATCTCTGGTCAGGATCTTAATTTGTCTCCAGCTCTG	1268
QY	5327	CTACCTCCCTTCCTGCGCGTTCCTCTCTCTGTCCAGCTGGGCACTGTGGTGGCTCGTT	5386
Db	1269	CTACCTCCCTTCCTGCGCGTTCCTCTCTCTGTCCAGCTGGGCACTGTGTGGCTCGTT	1328
QY	5387	CAGCTGTGTGTCCACATTTCTCAGGATTTCTCGAAAAGTTTAAACGAGTGTAGATTTTCCC	5446
Db	1329	CAGCTGTGTGTCCACATTTCTCAGGATTTCTCGAAAAGTTTAAACGAGTGTAGATTTTCCC	1388
QY	5447	CTGTAGACAGCAGATCACGATTTCTCCCGGAAAGTCAGGCTTCAGCCCTCTCTTTCTCTGC	5506
Db	1389	CTGTAGACAGCAGATCACGATTTCTCCCGGAAAGTCAGGCTTCAGCCCTCTCTTTCTCTGC	1448
QY	5507	CCAGCTGCGCGGCACTCTTAGCAAAACCTCAGGACCCCTTAACCCCAATAGACCTCTGACA	5566
Db	1449	CCAGCTGCGCGGCACTCTTAGCAAAACCTCAGGACCCCTTAACCCCAATAGACCTCTGACA	1508
QY	5567	GAGAAAGCAGGACATTTATCATGAGATCTCGTGTGGAGAGCCATATAGGCTTACGGTGTAAAGA	5626
Db	1509	GAGAAAGCAGGACATTTATCATGAGATCTCGTGTGGAGAGCCATATAGGCTTACGGTGTAAAGA	1568
QY	5627	GGCAGGGGAAGTGTGTGTGTAGAAAAGTCAGACCTTACATAGAAAGCTTAGCCACACAG	5686
Db	1569	GGCAGGGGAAGTGTGTGTGTGTAGAAAAGTCAGACCTTACATAGAAAGCTTAGCCACACAG	1628
QY	5687	AAAAGACAGACAGATCCCTCTATCTCCGCCCATTAAGATTTGAGT	5731
Db	1629	AAAAGACAGACAGATCCCTCTATCTCCGCCCATTAAGATTTGAGT	1673

RESULT 11	AX546496	1679 bp	DNA	linear	PAT 26-NOV-2002
LOCUS	AX546496				
DEFINITION	Sequence 5 from Patent WO0249669.				
ACCESSION	AX546496				
VERSION	AX546496.1	GI:25811663			
KEYWORDS					
SOURCE					
ORGANISM	Mus sp.				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciuromorphi; Muridae; Murinae; Mus.				
REFERENCE	1				
AUTHORS	Schu,C.A., Padua,R., Bonner,M., Donovan,M.G. and Soykan,O.				
TITLE	Electrically responsive promoter system				
JOURNAL	Patent: WO 0249669-A 5 27-JUN-2002; Medtronic, Inc. (US)				
FEATURES					
source	1..1679				
	Location/Qualifiers				
	1..1679				
	/organism="Mus sp."				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:10095"				
ORIGIN					
	Query Match	22.3%	Score 1279.8;	DB 6;	Length 1679;
	Best Local Similarity	82.2%;	Pred. No. 0;		
	Matches 1616;	Conservative 0;	Mismatches 57;	Indels 292;	Gaps 1;
QY	3767	GAATTCCTTCTCTCAAGGAAAGGAAAGTACGTCGACCTGCAAGAGTGATGCTCCCT	3826		
DB	1	GAATTCCTTCTCTCAAGGAAAGGAAAGTACGTCGACCTGCAAGAGTGATGCTCCCT	60		
QY	3827	AGACATCAAGACTTGTCTCTGGGGAGGCACACACTGTCGAACTTCAGTCTGAGAGACTA	3886		

Db	61	AGACATATATGACTTTTGTCTCTGGGGAGCCAGCACTGTGAACTTCAGTCTGAGAGAGTA	120
QY	3887	GGAAGCTCCCTCAGCCTGAAGCTATGCAATGACCAGGGTTGAAAGGGGAAAGGAGAG	3946
Db	121	GGAGGCTCCCTCAGCCTGAAGCTATGCAATGACCAGGGTTGAAAGGGGAAAGGAGAG	180
QY	3947	CTGGGATGGAGCTTGCTGTGTGGAGGCAAGGGGACGATATTAAAGCTGGAAAGAAAG	4006
Db	181	CTGGGATGGAGCTTGCTGTGTGGAGGCAAGGGGACGATATTAAAGCTGGAAAGAAAG	240
QY	4007	TGACCTTTACCAAGTTGTTCAACTCACCTTCAGATTAAAAATTAATCGAGTTAAGGCT	4066
Db	241	TGACCTTTACCAAGTTGTTCAACTCACCTTCAGATTAAAAATTAATCGAGTTAAGGCT	300
QY	4067	GGGTAGGGGAGGTGTGTGAGACGCTCTGTCTCTCTCTGCAATGCCCTGAGCCCTTGG	4126
Db	301	GGGTAGGGGAGGTGTGTGAGACGCTCTGTCTCTCTCTGCAATGCCCTGAGCCCTTGG	360
QY	4127	GGGAGGAGAAATGTGCCCAAGACTAAAAAAGGCTAGAGCCAGAGGGGCGAGGCA	4186
Db	361	GGGAGGAGAAATGTGCCCAAGACTAAAAAAGGCTAGAGCCAGAGGGGCGAGGCA	420
QY	4187	CAGACCTTTATGSGGCAAACTTGSGGGCCGTAAGTATGCAATGACAAAGAACTGGCAAT	4246
Db	421	CAGACCTTTATGSGGCAAACTTGSGGGCCGTAAGTATGCAATGACAAAGAACTGGCAAT	480
QY	4247	CGAATCCCTTCCTCTTAAAGGACGACGAGGAGGAACTCGAATTTTCACTCCCTATCATGTG	4306
Db	481	GGATCAAAAGGAGGAGGACGACGAGGAGGAAAGTGGAGGAG-----	525
QY	4307	ATAGAGAAAAGTGAAGTCAAGTTTCACTCCTATCAATGATAGAAAAAGTGAAGT	4366
Db	526	-----	525
QY	4367	CGAGTTTACCACTCCCTATCAGTATAGAGAAAAAGTGAAGTGAAGTTTCACTCCCTA	4426
Db	526	-----	525
QY	4427	TCAATGATAGAAAAAGTGAAGTCAAGTTTCACTCCTATCAGTATAGAAAAAGT	4486
Db	526	-----	525
QY	4487	GAAAGTCAGATTTACCACTCCCTATCAGTATAGAAAAAGTGAAGTGAAGTTTACAC	4546
Db	526	-----	525
QY	4547	TCCTTATCAGTATAGAAAAAGTGAAGTCAAGCTCGTATCCAGCAGAGACTTCCAAT	4606
Db	526	-----GGTCCGACAGAGGACTTCCAAT	548
QY	4607	TTAGGAGCAGGCACTATGGATGGATATTAAGGGGCTGAGGACATGAGCTGTCAAG	4666
Db	549	TTAGGAGCAGGCACTATGGATGGATATTAAGGGGCTGAGGACATGAGCTGTCAAG	608
QY	4667	ATTTCCTCAACCCAGGTAGAGGAGTTTGGGTGGGGCTTTCAACCAACACAGACT	4726
Db	609	ATTTCCTCAACCCAGGTAGAGGAGTTTGGGTGGGGCTTTCAACCAACACAGACT	668
QY	4727	CTCCCACTTAAGAAAGAAATGCTCTTCTGGAAGTGGGGTTCAAGGCCGTCAAGATCT	4786
Db	669	CTCCCACTTAAGAAAGAAATGCTCTTCTGGAAGTGGGGTTCAAGGCCGTCAAGATCT	728
QY	4787	GACAGGGGTGGCTTCCACAGCCTGGGAAAGTTCAGTGGGACGAGGTTTCAAGAA	4846
Db	729	GACAGGGGTGGCTTCCACAGCCTGGGAAAGTTCAGTGGGACGAGGTTTCAAGAA	788
QY	4847	CACATGATGCCCCCTTCCCTTACGTGTCTTCAATCTTCTCTGGGAGTCTCTCC	4906
Db	789	CACATGATGCCCCCTTCCCTTACGTGTCTTCAATCTTCTCTGGGAGTCTCTCC	848
QY	4907	CGTCTTGATTATTTGGCTTGTGTCTTCAAGCAAGTTTGCCTGTGTCTTCACTTCA	4966

Db 849 CGTCTGGTTTATCTGGCTCTTCGTCTGAGCAAGATTTGGCCCTGTGCTGCATCTCA 908
 Qy 4967 TCTTTCTCTACTGCTCCGATGCTTGCCTTGTGGGTGCTTCTCTTCCATCCCA 5026
 Db 909 TCTTTCTCTACTGCTCCGATGCTTGCCTTGTGGGTGCTTCTCTTCCATCCCA 968
 Qy 5027 TTTCTCACTTCACTTTTCTCCCTTCTCAATTGTATTCATTCCTTCTCTTCTCT 5086
 Db 969 TTTCTCACTTCACTTTTCTCCCTTCTCAATTGTATTCATTCCTTCTCTTCTCT 1028
 Qy 5087 TCT 5146
 Db 1029 TCT 1088
 Qy 5147 TCT 5206
 Db 1089 TCT 1148
 Qy 5207 TAAACAACTCTTCAGATGAGCCACAGCTTCAGTGTGCTGCTCTTAACTCTCTCA 5266
 Db 1149 TAAACAACTCTTCAGATGAGCCACAGCTTCAGTGTGCTGCTCTTAACTCTCTCA 1208
 Qy 5267 CCCCCTGGCTTGTCT 5326
 Db 1209 CCCCCTGGCTTGTCT 1268
 Qy 5327 CTACT 5386
 Db 1269 CTACT 1328
 Qy 5387 CAGCTGTGTCTCACATTTCTTCAAGATTCTCTGAAAGTTTAAACAGTGAAGTTTCCC 5446
 Db 1329 CAGCTGTGTCTCACATTTCTTCAAGATTCTCTGAAAGTTTAAACAGTGAAGTTTCCC 1388
 Qy 5447 CTGTAGACAGAGATTCAGATTTCTCCCGGAAGTACAGCTTCCAGCTCTCTCTCTCTGC 5506
 Db 1389 CTGTAGACAGAGATTCAGATTTCTCCCGGAAGTACAGCTTCCAGCTCTCTCTCTCTGC 1448
 Qy 5507 CCAGCTGCCCGGCACTCTTACCAAACTTACAGGCACTTACCCGACATAGACTCTGACA 5566
 Db 1449 CCAGCTGCCCGGCACTCTTACCAAACTTACAGGCACTTACCCGACATAGACTCTGACA 1508
 Qy 5567 GAGAAGCAGAGCATTTTACATGAGTCTGTGTGAGAGGCACTTACCGTGTAAAGA 5626
 Db 1509 GAGAAGCAGAGCATTTTACATGAGTCTGTGTGAGAGGCACTTACCGTGTAAAGA 1568
 Qy 5627 GGCAGGGAAGTGTGTGTAGAGAAATCAGACATTTCAATAGAGCTTACGCCACACAG 5686
 Db 1569 GGCAGGGAAGTGTGTGTAGAGAAATCAGACATTTCAATAGAGCTTACGCCACACAG 1628
 Qy 5687 AATGACAGACAGATCCCTCTATCTCCCCATTAAGATTGAGT 5731
 Db 1629 AATGACAGACAGATCCCTCTATCTCCCCATTAAGATTGAGT 1673
 RESULT 12
 MUSCHHG 2594 bp DNA linear ROD 27-Apr-1993
 LOCUS Mus musculus cardiac myosin heavy chain gene, 5' end.
 DEFINITION M62404.1 GI:192609
 ACCESSION M62404.1
 VERSION M62404.1
 KEYWORDS cardiac myosin heavy chain.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2594)
 AUTHORS Gulick,J., Subramaniam,A., Neumann,J. and Robbins,J.
 TITLE Isolation and characterization of the mouse cardiac myosin heavy
 chain genes
 J. Biol. Chem. 266 (14), 9180-9185 (1991)
 JOURNAL 9125025
 MEDLINE 2026617
 PUBMED

COMMENT Original source text: Mus musculus DNA.
 FEATURES Location/Qualifiers
 source 1..2594
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 mRNa join(603..623,1305..1373,1665..1880)
 /product="cardiac myosin heavy chain"
 exon 603..623
 /number=1
 intron 624..1304
 /number=1
 exon 1305..1373
 /number=2
 intron 1374..1664
 /number=2
 exon 1665..1880
 /number=3
 CDS 1680..1880
 /codon_start=1
 /product="cardiac myosin heavy chain"
 /protein_id="AA37424.1"
 /db_xref="GI:192610"
 /translation="MTDAQMDPGAAAGTLRSSEKRELEAQTREPPDIRTECFVDDKE
 EYKAKVVSREGGKYATETENGK"
 1881..>2594
 /number=3
 intron
 ORIGIN
 Query Match 22.3%; Score 1279.8; DB 10; Length 2594;
 Best Local Similarity 82.2%; Pred. No. 0;
 Matches 1616; Conservative 0; Mismatches 57; Indels 292; Gaps 1;
 Qy 3767 GAATTCCTTACTATCAAGGGAAGTGAAGTGTGACCTGCAAGTGAATGCTCTCCT 3826
 Db 1 GAATTCCTTACTATCAAGGGAAGTGAAGTGTGACCTGCAAGTGAATGCTCTCCT 60
 Qy 3827 AGACATCATGACTTTGTCTGTGGGAGCCAGCACTGTGAACTTCAAGTCTGAGAGTA 3886
 Db 61 AGACATCATGACTTTGTCTGTGGGAGCCAGCACTGTGAACTTCAAGTCTGAGAGTA 120
 Qy 3887 GAGAGCTCCCTCTAGCCTGAAGCTATGCAATAGCCAGGTTGAAAGGGGAAAGGAG 3946
 Db 121 GAGAGCTCCCTCTAGCCTGAAGCTATGCAATAGCCAGGTTGAAAGGGGAAAGGAG 180
 Qy 3947 CCGGAGTGGAGCTGTGTGTGTGAGAGCGAGGGAAGATTAAGCTTGAAGAGAG 4006
 Db 181 CCGGAGTGGAGCTGTGTGTGTGAGAGCGAGGGAAGATTAAGCTTGAAGAGAG 240
 Qy 4007 TGACCCCTTACCCAGTTGTCACTCAACCTTCAAGATTAATAATTAAGCTTAAGGCT 4066
 Db 241 TGACCCCTTACCCAGTTGTGTCACTCAACCTTCAAGATTAATAATTAAGCTTAAGGCT 300
 Qy 4067 GGGTAGGGAGGTGTGTGTGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4126
 Db 301 GGGTAGGGAGGTGTGTGTGAGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 360
 Qy 4127 GGGAGGAGGAATGTGCCCAAGGACTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4186
 Db 361 GGGAGGAGGAATGTGCCCAAGGACTTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
 Qy 4187 CAGACCTTTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4246
 Db 421 CAGACCTTTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
 Qy 4247 CGATACCTTTCTCTTCTTAACGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 4306
 Db 481 CGATACCTTTCTCTTCTTAACGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 525
 Qy 4307 ATAGAGAAAGTGAAGTGAAGTTTCACTCCCTATCAAGTGAATGAGAAAGTGAAGT 4366
 Db 526 ----- 525


```

-----
Percentage of bases with a quality value >= 40 : 99 %.
Location/Qualifiers
1..157910
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone="C-2201G16"
/clone_lib="Caltech-D"
13550..13694
/notes="matching EMBL:AA452257
RHdb:RH92419
dbSTS:STS65422
Identified using the e-PCR software (G. Schuler)"
46758..46857
/notes="matching EMBL:N91549
RHdb:RH76320
dbSTS:STS3407
Identified using the e-PCR software (G. Schuler)"
57409..57527
/notes="matching EMBL:X52889
RHdb:RH13799
dbSTS:STS7881
Identified using the e-PCR software (G. Schuler)"

STS

ORIGIN

Query Match      9.7%: Score 558.2; DB 9; Length 157910;
Best Local Similarity 63.7%: Pred. No. 3,4e-164;
Matches 1380; Conservative 0; Mismatches 658; Indels 127; Gaps 30;

Qy 2096 CTCACCTGTGGCCAGAGCGTGCATCTGTGTCAACATCTCTAGAAATGTCAGACTGC 2155
Db 44522 CTCACCTGTGGCCAGAGCGTGCATCTGTGTCAACATCTCTAGAAATGTCAGACTGC 44463

Qy 2156 AGGGCT-----GCGTTGGAGAGCGCTGGAAAAGATGTGTGAGAGCCAGGGAGACA 2207
Db 44462 GGAAGCTGTGTGCAAGCGCTTAGGGAGAGCGGAGACCAATATGTGTGAGAGCGAGAGAGA 44403

Qy 2208 AGGGGGCTTAGAAAAGAGAGAGGGCAAAACAGGCCACACAAGAGGGAGAGCCAGA 2267
Db 44402 A-GGAGCGCTGGAGAGAGAGAGCGGTACAGCGAGACAGC-----GGGAGGGCGCTGGG 44348

Qy 2268 ACTGAGTTAACTCCCTTCCTTGTTCATCTTCATAGGAGGAGCGAGACTCTGTGTGACA 2327
Db 44347 GCTGAGTGAACACTTCCTTCCTTCCTTCACAGAGAGGGGGACTACCCCCACACAC 44288

Qy 2328 CCATCCCCCATG-----AGCCCCCATCACTCCATACCAAGTTTGGCTGTGAGTGCATTCT 2381
Db 44287 CTCACACTCCCCGCTGCCCTGTCCACACCACTATACCAAGTGAAGGCGCTGTCTCAGAGACA 44228

Qy 2382 AGGTTCCCTGAGAGCAGAGCGCTGTCCTTGTCTTGTGACCTGACCCAGGTGACCCCAAT 2441
Db 44227 GGGTTCCTGTAGAGCAAAAGCTGTGATCTTTGTCACTGACCCGACCCAGACTGACTGTGC 44168

Qy 2442 GTTCTCAGTACTCTTGTGATGCTTCAAGAGCTTGAGAACAGGAGTGAACATATTAGGCC 2501
Db 44167 GTTCTCAGTACTCTTATATATGTCCCAAGCTATGAGAACCAAGAGTGAAGCACT-GGCC 44109

Qy 2502 ATGGGCTTAACCTCTGAGCTTGACACAGAGAGCTCAAGTGAAGCTTCCAGGAGACAGCTGC 2561
Db 44108 ATGGGCTTAACCTCCCGCTCAAAATTCGTGGGCTCTGAGTAACCT-----AGAGCTGG 44057

Qy 2562 AGACAGGAGGCTTTATCCCAAGAGCAACATTTGGCATAGGTGGCTGC---AAATGG 2618
Db 44056 GGAAGAGGTGGCTGTGATCTCCCTGGGGGTATATGATTTGGCATTAAGCACTGCAAGAAATGG 43997

Qy 2619 GAATGCAAGGTTGATCAAGTCTCCCTTCAAGATATCTGATGACAGACCTTAAGACCCCTGG 2678
Db 43996 GGAATGTGGGGTCAACTAAGTCTCCCTCAAGACATTTG-AGAAAAAGTCTCAAAATGCTCTGG 43938

Qy 2679 AGAAGAGGGGTATGCTCTGCTCCCAACCAACCAATAAGGGAAGTGAATATCTTAAGGGGCT 2738

```

```

Db 43937 AAAGAGAGGTGTGTTCTCGAGCCCA-----GGAATGGGTGGAGCGCTCTAAGGAGCT 43884
Qy 2739 GCGGACCTTGGGGAGACACCACTTACTGAGAGTGTGAGCCAGAAAAATGACCGGCC 2798
Db 43883 AAGG-----GAGAGGCCACATCTTATGGAAGAGTGTCTCTACCAAGAGAGTGAACCTCT 43829
Qy 2799 TGTGTCTTGGCCCACTTCCACACTTAAGAGTATATTAGAGGTGACAGTATAGGTGG 2858
Db 43828 GTGTCTTACCCACCTCCACACCTTGAAGCTATATTGAGAGGTGACAGTAACTGTGTGG 43769
Qy 2859 GAGCTGTGACAGGAGAGGTGTCTCGGTGTGAGGGGTGAGGGGAAAGCCAGAGAGGG 2918
Db 43768 ---AGGGGGCTGGAGAGGTGTCTCGGTGTGAGGGTGTG-GGGGAGAGGCCAGAGAGGG 43713
Qy 2919 GAGTGTGCTTGTGTCTCTGAAACACAATGTCTACTTATTAGTTATACAGCATGACTGCTA 2978
Db 43712 GAGTGTGCTTGTGTGTTCGAAACACAATGTCCACTTATGATTAACAGAAAGAGCTGTGCT 43653
Qy 2979 AAGACCCAACTATACAGACTCTTGAAGAAGACAGAGCCCTGGAGGA---CAGGGTGT 3034
Db 43652 AAGACCTGACACTTACAGCTCTG-AGAGATGGCAGCATTTGAGAGATGGGGAGCGGGGCC 43594
Qy 3035 CTCTGAGCGCTTGGGTGCTTGAATGTGCGCACAAAAGAGGCGATGAGTGAATAAAGCC 3094
Db 43593 TTGAGGGGCTTGAAGGTGTGTGAGAGACACCGCAAGAGG---GAATGTAGATTAAGGCC 43537
Qy 3095 CCAGAGC-----GTTAGAGAAAGGCACTTGGAAAGGGGTCACTGTGAGAGGCC 3144
Db 43536 TGGAGGGCTTTGGAAAAGGGCAACAAATTTCTGCTGGGAGAGGGGCGAGGTGGCAGAGGCC 43477
Qy 3145 CTATCC-ATGAAATCTGAGAGCTGTGGG-CCAATGTGTAAATCTTGTGGCGCTGCCAGGC 3202
Db 43476 CACCCGATGGAATGACAGTGTGCGGAGACCAATGTGTAAATCCCAAGCGCTGCAG-- 43419
Qy 3203 ATTCAAGAGACACCTGCATCTCTGTGAGCGCTGTGGGAGAGCGGAAAGAGAACCCCA 3262
Db 43418 -----GACCTGCACCTCTGTGAGCCCGGTGTGAGAGAGAGAGAGAGAGCCCTTCA 43369
Qy 3263 CTATATCCCTTCTCTCCCTAGCCGCCAGATTAAACACTGTGAGCTTCCCTTCCGACCT 3322
Db 43368 CCAACCCCTCTCTCTCAGGCGCTAAGATTAGCAGCATCTTCTT-----CTGCCACCC 43313
Qy 3323 CCCATGAGAGTGTGAGAGGTGTTCAGAGAGGGGTAAACCTTACTATGTCCAAACATCATGG 3382
Db 43312 CACCCGAGAGACAGAGGTGTGAG-GGGAGGGCAAAAGCTATGTGTGCAAAATGCGGATG 43254
Qy 3383 TGCAGATATATGATCAGATGTGTGAGAGCA---AGAAAGGAATCTGACAGGCTTAAC 3439
Db 43253 TGCA-ATGTGTGATCTCATGTACACAAAAGAAAGTGAAGAAATCTGCAAGCTTGAAG 43195
Qy 3440 TGGGTTATGTGTAAAGTCTGTGTGATGTGTGTGTGTGACTGAAAACGGGCATGGCT 3499
Db 43194 TGGGTTAAGTGTAAATCTGTGATGTGTGTGCTGAAGATGGGACAGACAGCG----- 43143
Qy 3500 GTGCACTGTTCAGTCTGTGTGTGAGTTTACAGACTGCAAGTTTGTGTAAATTTGCC 3559
Db 43142 -----TCAAGTGTGTATGTGAGAGTGTGAACTGGAGTTGTGTGAAAATCTGC 43093
Qy 3560 CAAGCAAGATGGGTGAATCCCTTCATGGTTTAAAGATTTGATGATGAGCTGCATCT 3619
Db 43092 CTGAGGGGCGAGGAGAAATCATCTGCAATTCGTGAGACAGTTTGAATGTGGCACTCTAT 43033
Qy 3620 CAAGCAATGTGAAAAATGAAATGAGACACTTATATGTGTCTTAAGCTAAGGTAG---CA 3676
Db 43032 C-AGGAGCATTTAGGAAAGGGGTGGGAGCTTCAAGCGTGTCCCAAAACAGAGGTGGCTCA 42974
Qy 3677 AGGCTTTTGGAGACACTGTCTTGAAGTGTGGGCAACAGAGCTACAGACATATCTGT 3736
Db 42973 AGACTTTGGAGAACTTGTCTGAAGACTTGGGGAACAG-----AAGGAGACAGGCAT 42919
Qy 3737 ACAGAGTAAAGAGAGAGAGAGGGGGGTGAATTTCTTACTATCAAAAGGAACTGAG 3796

```

```

Db      42918 GGCATTATGACAGTGAAGCCAGACAGA-----ATTCTGACAAAAAAACTGAG 42865
Qy      3797 TCGTGACCTCGAAAGTGAATGCTCTCCCTAGACATCATGACTTGTCTCTGGGGAGCCA 3856
Db      42864 CCAATGGAGATGACAACAGATCCCTTCCCTGGGACCAT-ACAGCAGCTTTTATGCTCTA 42806
Qy      3857 GCACGTGGAACCTTCAGGTCTGAGAGTGAAGAGCTCCCTCAGCCTGAAGCTATGACAG 3916
Db      42805 GCACGTGGGGCTCCAGACTACTAACAAGAGAAATGCTCCAGCCTGGAGCTGTGTAGAGG 42746
Qy      3917 ATAGCCAGGTTAAAGGGGAAAGGAGAGCTGGAGTGGAGCTTGTGTGTGAGGCA 3976
Db      42745 AGGTCAAGATGGAGAGAGAGCTGGGGAACAGGGAGAAAGCCATGGTTGGAGAGG 42686
Qy      3977 GGGGACAGATATTAAGCTGGAGAGAGAGTGAACCTTACCAGCTTCAATCAGCCCT 4036
Db      42685 GAGGACAGGCAATTTGGCTGACAGAGAGAGTGAACCTCACCAGTGTTCAGTTCACCCCT 42626
Qy      4037 TCAGATTTAAATACTAGAGTAAGGCG-----CTGGGTAGGGAGGTGTGTGAGAGCG 4091
Db      42625 TCGGTTTAAATACTAGAGTAAAGGCCATGCGAGAGGTGGAGAGCGGTGTGAGAGAG 42566
Qy      4092 TCGTGTCTCTCTCTGTCGATGAGCCCTTGGGAGAGAGAAATGTGCCCAAGACT 4151
Db      42565 TCGTGTCTCTCTCTGTCGATGAGCCCTTGGGAGAGAGAAATGTGCCCAAGACT 42506
Qy      4152 AAAAAAGGCGATGAGAGCCGAGAGGGCGAGAGCGCAAGACTTTCATGGGCAAACTTGG 4211
Db      42505 AAAAAAGGCGCGTGGAGCCAGAGAGGCTGGGAGAGAGAACCTTCTTGGGCAAACTGAG 42446
Qy      4212 GGCCC 4216
Db      42445 GGCCC 42441

RESULT 14
CNS0000B/c 196292 bp DNA linear PRI 22-MAY-2001
LOCUS      Human chromosome 14 DNA sequence BAC R-124D2 of library RRC1-11
DEFINITION From chromosome 14 of Homo sapiens (Human), complete sequence.
ACCESSION AL049829
VERSION    AL049829.4 GI:8217859
KEYWORDS   HTG.
SOURCE      Homo sapiens (human)
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 196292)
AUTHORS    Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
            Brothier,P., Catolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
            Levy,M., Beckenberg,R., Bruls,T., deBerardinis,V., Cruaud,C.,
            Gyapay,G., Saurin,W. and Weissenbach,J.
            Sequencing of the human chromosome 14
            Unpublished
            2 (bases 1 to 196292)
TITLE      Genoscope.
JOURNAL    Direct Submission
AUTHORS    Submitted (21-MAY-2001) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT    On Jun 3, 2000 this sequence version replaced gi:6138746.

----- Genoscope / Centre National de Sequencage
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: SeqRef@genoscope.cns.fr

-----
The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : R-244B17
Downstream BAC (overlapping the SP6 end) : C-2201G16 (AC=AL132855)
----- Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 7.94x in Q20 bases; sum-of-contigs

```

```

-----
Overall quality chart :
Range : bases
0 - 9 :
10 - 19 :
20 - 29 :
30 - 39 :
40 - 49 :
50 - 59 :
60 - 69 :
70 - 79 :
80 - 89 :
90 - 99 :

Percentage of bases with a quality value >= 40 : 99 %.

-----
FEATURES
source
1..196292
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="caxon:9606"
/chromosome="14"
/clone="R-124D2"
/clone_1b="RRC1-11"
78259..78398
/note="matching EMBL:AA908790
RHdb:RH102162
dbSTS:STS69699
Identified using the e-PCR software (G. Schuler)"
80433..80566
/note="matching EMBL:AA167748
RHdb:RH8727
dbSTS:STS68485
Identified using the e-PCR software (G. Schuler)"
82164..82299
/note="matching EMBL:H79035
RHdb:RH78773
dbSTS:STS5514
Identified using the e-PCR software (G. Schuler)"
93580..93808
/note="matching EMBL:AA007328
RHdb:RH75103
dbSTS:STS52191
Identified using the e-PCR software (G. Schuler)"
107758..107951
/note="matching EMBL:R94929
RHdb:RH65111
dbSTS:STS45044
Identified using the e-PCR software (G. Schuler)"
107996..108181
/note="matching EMBL:D11677
RHdb:RH47139
dbSTS:STS40201
Identified using the e-PCR software (G. Schuler)"
128387..128525
/note="matching EMBL:H2648
RHdb:RH53698
dbSTS:STS18321
Identified using the e-PCR software (G. Schuler)"
138462..138738
/note="matching EMBL:R59134
RHdb:RH53972
dbSTS:STS42930
Identified using the e-PCR software (G. Schuler)"
138509..138658
/note="matching EMBL:M78864
RHdb:RH95543

```



```

STS      RHdb:RH1243
          dbSTS:STS24710
          Identified using the e-PCR software (G. Schuler)
          /note=.138677
          /note="matching EMBL:M78946
          RHdb:RH28671
          dbSTS:STS20163
          Identified using the e-PCR software (G. Schuler)
          /note=.138647
          /note="matching EMBL:M78986
          RHdb:RH28416
          dbSTS:STS20162
          Identified using the e-PCR software (G. Schuler)
          /note=.138926
          /note="matching EMBL:H72023
          RHdb:RH68557
          dbSTS:STS48564
          Identified using the e-PCR software (G. Schuler)
          /note=.161761
          /note="matching EMBL:AA452257
          RHdb:RH92419
          dbSTS:STS65422
          Identified using the e-PCR software (G. Schuler)
          /note=.194903
          /note="matching EMBL:N91549
          RHdb:RH76320
          dbSTS:STS53407
          Identified using the e-PCR software (G. Schuler)

STS      9.7%; Score 558.2; DB 9; Length 196292;
Query Match Best Local Similarity 63.7%; Pred. No. 3.6e-164;
Matches 1380; Conservative 0; Mismatches 658; Indels 127; Gaps 30;

ORIGIN
2096 CTCACCTGTGGCCGAGAGGTCATCTGTGTCACATCTCTAGAATGTCATGACTGC 2155
    |||||
192568 CTCACCTGTGGCCGAGAGGTCATCTGTGTCACATCTCTAGAATGTCATGACTGC 192509
    |||||
2156 AGGGCT-----GGCTTGGAGGAGAGCTGGAAGATGTGAGAGCCAGGGAGACA 2207
    |||||
192508 GGAAGCTGTGTGACGAGGCTTGAAGAGAGGAGAGCAATATGTGAGAGGAGAGAGA 192449
    |||||
2208 AGGGGGCTTAGAAGAGAGAGGAGAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAG 2267
    |||||
192448 A-GGAGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 192394
    |||||
2268 ACTGAGTAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2327
    |||||
192393 GCTGAGTGAACACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 192334
    |||||
2328 CCATCCCCCATG-----AGCCCCCATGCCCATACCAAGTTTGGCTGAGTGCATTCT 2381
    |||||
192333 CTCACGCTCCCGGCGTGCCTGACACACCACTATACCAAGTGAAGGCTGCTCAGAGACA 192274
    |||||
2382 AGGTTCCCTGAGAGACAGAGCTGGCTTTGTCTTGTGACCTGACCCAGTGAACCAAT 2441
    |||||
192273 GGGTTCCCTGAGAGACAAAGCTGGTCTTGTGACCTGACCCGAGCCAGAGTGAAGCTGGC 192214
    |||||
2442 GTTCTAGTACCTTTGAATGCGCTCAAGAGCTTGAAGAGAGAGAGAGAGAGAGAGAGAG 2501
    |||||
192131 GTTCTAGTACCTTTGAATGCGCTCAAGAGCTTGAAGAGAGAGAGAGAGAGAGAGAG 192155
    |||||
2502 ATGGGCTTAACCTCTGAGAGCTTGCACACAGAGAGCTTCAAGTGAAGCTTCAAGGAGACAGAGCTGC 2561
    |||||
192154 ATGGGCTTAACCTCCCGGCTCAATTTGGGCTCTGAGTAACT-----AAGAGCTGG 192103
    |||||
2562 AGAAGAGTGGCTTTATTCCTCAAGAGAGACATTTTGGCATAGTGGCTGC--AAATGG 2618
    |||||
192102 GGAAGAGTGGCTCTCATTCCTGGGGGTAATGATTTGGCATTAACACCTGCAAGAGATGG 192043
    |||||
2619 GAATGCAAGGTTGATCAGTCCCTTCAAGATATCTGATGACAGAGACTTAAGACCTCTGG 2678
    |||||
192042 GGAATGGAGTCACTAAGTTCCTTCAAGAGACATTG-AGAAAGGTCCTCAATGCTCTGG 191984
    |||||

```

```

2679 AGAAGAGGCTATGCTCTGCCCCACCCACCATTAAGGAGAGTGAACCTATCTTAGGGGCT 2738
    |||||
191983 AAGAGAGGCTATGCTCTGAGCCCA-----GGAATGGGGTGGCAGCGTCTTAGGGAGCT 191930
    |||||
2739 GCGGACCTTGGGAGAGACCACTTACTGAGAGTGTGAGCCAGAGAAAAGTACGCGGCC 2798
    |||||
191929 AGGG-----GAGAGGCGACATCTTAATGGAAGAGTGTCTTACAGAGAGCTGACCTCT 191875
    |||||
2799 TGTGCTCTGCCCCACCTCCACACTCTAGAGCTATATTAAGAGGTGACAGTATAGAGGTGG 2858
    |||||
191874 GTGCTCTTACCCACCTCCACACCTCTGAGCTATATTAAGAGGTGACAGTAACTGGTGGG 191815
    |||||
2859 GAGCTGTAGCAGAGAGAGAGTCTTCTGGGTGTGAGGCTGTAGGGGAAAGCCAGAGCAGGG 2918
    |||||
191814 ---AGGGGGCTGGAGGGGTGTCTTGGGTGTGAGGGGTG-GGGGGAAGGCGAGAGCAGGG 191759
    |||||
2919 GAGTGTGCTTTGTCTCCGGAACAGATGCTTAAGTTATTAAGGATGACCTGCTTA 2978
    |||||
191758 GAGTGTGCTTTGTCTCCGGAACAGATGCTTAAGTTATTAAGGATGACCTGCTTA 191699
    |||||
2979 AAGACCAACATCTAGACCTCTGAAAAGACAGACAGCCCTGAGAGA-----CAGGGGTTGG 3034
    |||||
191698 AAGACCTGACCTTACAGCTCTG-AGAGATGACAGCATTTGAGAGATGGGAGGCGGGCC 191640
    |||||
3035 CTCTGAGCCTTGGGTCTTGTATGTGTCCCAAAAGAGGCGATGATGTGATTAAGGCC 3094
    |||||
191639 TTGAGGCGCTAGGGTCTTGTAGAGACACCAAAAGAGG---GAATGTGATTAAGGCC 191583
    |||||
3095 CCAAGAGC-----GTTAGAGAGGCACTTGGGAAAGGGGTGAGTCTGACAGGCC 3144
    |||||
191582 TGGAGGCGCTTGGAAAGGCGCAAAATTTCTGTGGGAGGGGGGAGGTGGAGAGCCC 191533
    |||||
3145 CTATCC-ATGGAATCTGAGAGCTGTGGG-CCAACTGTGTAAATCTGTGGGCTGCCAGGC 3202
    |||||
191522 CACCCGATGGAATGAGAGTGCCTGGGAGCCAGATGTGTAAATCCCAAGGCTGCCAG-- 191465
    |||||
3203 ATTCAAGACAGACCTGTATCTCTGGCAGAGCTGGGAGGCGGAGAGGAGCAACCCCA 3262
    |||||
191464 -----GCACCTGACACCTCTGTGGCAGGCGGTGAGGAGAGAGGAGCAACCCCTCA 191415
    |||||
3263 CTATACCCCTTCTCCTCAGAGCCCGAGGATTAACCTCTGCGCTTCCCCCACT 3322
    |||||
191414 CCAACCCCTCTCTCTCTAGAGCCCTTAGATTAAGACACATCTTCTT---CTGCCAACC 191359
    |||||
3323 CCATCAGAGAGTGAAGGTTGTCAGAGAGAGGTTAAACCTACATGTCCAAACATCATGG 3382
    |||||
191358 CACCCAGAGAGAGAGAGTGAAG-GGAGGGCAAAAGCTATGTGTGCAATGGGATG 191300
    |||||
3383 TGCAGATATATGATCATGATGTGTGAAGGCA--AGAAAGAAATCTGCAGGCTTTAC 3439
    |||||
191299 TGCA-AATGTGTATCTCAATGTACAAAAAGAAAGTGAAGAAATCTGCAGGCTTTAG 191241
    |||||
3440 TGGGTTAATGTGAAGTGTGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3499
    |||||
191240 TGGGTTAATGTGAAGTGTGTGTGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 191189
    |||||
3500 GTGACGCTGTTCAGTTCTGTGCTGTGAGGTTTACAGACTGACAGGTTTGTGTAAATGGC 3559
    |||||
191188 -----TCAAGTGTGATGTGAAGTGTGAACTGGAGTTTGTGTGAAGAAATCTGC 191139
    |||||
3560 CAAGGCAAAAGTGGTGAATCCCTTCCATGTGTTTAAAGATTTGATGTGTGCTGCATCT 3619
    |||||
191138 CTGAGCGGAGGAGAAATACATGCGCATTCGAGACAGAGTGTGAATGTGGCCACTAT 191079
    |||||
3620 CAAGGCAATGAAATTAAGATGAGACACTTATATGTGTCTTAAGCTAAGTAG--CA 3676
    |||||
191078 C-AGAGCATTTAGGAGAGAGGAGGAGATCTCAACATGTGTCCAAACAGAGGTGGCTCA 191020
    |||||
3677 AGGCTTTTGGAGAGACCTGTCTTAGATATGTGGCAACAGAGACTACAGACATATCTGT 3736
    |||||
191019 AGACCTTTGGAGAGACCTGTGTCTGAAGACTTGGGAAACAG-----AAGAGAGCAAGCAT 190965
    |||||

```


[illegible]

exon 9308..9406
/number=10
exon 9954..10057
/number=11
exon 10150..10288
/number=12
exon 11771..12039
/number=13
exon 12322..12492
/number=14
exon 13709..14018
/number=15
exon 15118..15188
/number=16
exon 15467..15554
/number=17
exon 15644..15761
/number=18
exon 15906..16029
/number=19
exon 16303..16439
/number=20
exon 18400..18655
/number=21
exon 18815..19057
/number=22
exon 19205..19381
/number=23
exon 19666..19811
/number=24
exon 20071..20161
/number=25
exon 22275..22664
/number=26
exon 23001..23127
/number=27
exon 23209..23327
/number=28
exon 23663..23859
/number=29
exon 24380..24563
/number=30
exon 24791..24956
/number=31
exon 25061..25185
/number=32
exon 26079..26387
/number=33
exon 26570..26773
/number=34
exon 27658..27783
/number=35
exon 27990..28265
/number=36
exon 29313..29408
/number=37
exon 30071..30205
/number=38
exon 30569..30643
/number=39

ORIGIN

Query Match 8.6%; Score 494.8; DB 9; Length 31462;
Best Local Similarity 62.2%; Pred. No. 3.2e-144;
Matches 1340; Conservative 0; Mismatches 702; Indels 114; Gaps 31;

QY 2096 CTCACCTGTGGCAGAGGCTTCATCTGTCCACATACTTAGATGTTTCATCAGACTGC 2155
DB 2256 CTCACCTGTGGCAGAGGCTTCATCTGTCCATTAATTCAGAGCTCCCGACAGAAATGA 2315
QY 2156 AGGGCTGCTTGGAGGAGCTTGAAGAAAGATATGTGAGGACGAGGAGACAAAGGGGCC 2215
DB 2316 GGAAGCTGTG-TGCAAGGCAAGCTTAAGGAGAGCAGAGCAATATGTGTAGAGCAGGAGAG 2374

QY 2216 TAGAAAG-GAAGAAAGGCAAAACGAGCCACACAAGAGGGGAGGCCGAGACTGAGT 2274
DB 2375 AAGGAGCCTGGGAAAGGAGAGGACGGTACAGCGGAGACGGGCAAGGGCTGAGT 2434
QY 2275 TAACTCTCTCTTGTGTGATCTTCCATPAGAGGCAGTGGGAATCTGTGACCACTATCC 2334
DB 2435 GAACACCTTCTCTCTCTCTCTCCACAGAGGGGGGAGCTACCCCACTCCACTC 2494
QY 2335 CCATGAGCCC-----CCATACCCATACCAAGTTTGGCTGAGTGGCAATTCTAGTTCCC 2389
DB 2495 CCGCTGCGCTTGGCCACACCACTTATACAAAGTGAAGGCTGTCTCAAGACACAGGTTCCC 2554
QY 2390 TGAAGACAGAGCTGAGCTTTGTCTCTTGAAGCTGACCCAAAGTGAACCAATGTTCTAG 2449
DB 2555 TGAAGACAAAGCTGAGCTTTGTCACTGACCCCAAGCTGACCTGAGCTGCTCTCAG 2614
QY 2450 TACCTTGAATGCCCTCCAAAGAGCTTGAGAAACAGGCAGTGACATATTAGGCAATGGGCTA 2509
DB 2615 CTCCTTATCATGTCCCAAGAGCTATGGAACCAAGCAGTGAACGCACT-GGCCATGGGCTA 2673
QY 2510 ACCCTGAGCTTGACACAGAGAGCTCAAGTGACTCCAGGACACAGCTGACAGAGT 2569
DB 2674 ACCCCCCGCTGGAATTCGAGGAGCTGAGTAACCT-----AGAGCTGGGAGCAGT 2725
QY 2570 GGCCTTATCCCAAGAGCAACCATTTGGCATAGGTGCTGC---AAATGGAAATGCA 2626
DB 2726 GGCCTGATCCCG-TGGGGTAATGATTTGGCATTAACACCTGCAAGAAAGATGGGAGTGG 2784
QY 2627 GGTGAATCAGAGTCCCTTCAAGAAATCTGATGCAAGACTTAAGACCCCTGAGAGAGG 2686
DB 2785 GGTCAACTAAGGTCCCTCCAAAGACATTTG-AGAAAAGGTCCAATGTCTCGAGAAAGAG 2843
QY 2687 GTATGCTCTCGGCCCCACCAACCATTAAGGGAGTGAATCTATCTAGGGGGCTGGCACT 2746
DB 2844 GTGTGTTCTGAGCCCA-----GGAATGGGTGGGAGGCTCTTAGGAACTTAGG---- 2894
QY 2747 TGGGAGACACCACTTACTGAGAGTGTGAGGCCACAGAAAACTGACCGGCTGTGTCT 2806
DB 2895 -GAAGGCCACATCTTAATGGAAGAGTGTCTCTCCAGAGAACTGACCCCTGTGCTCT 2953
QY 2807 GCCCACTCCACACTTGAAGCTATATTGAGAGTGAAGTGAATAGAGGTGGAGTGTGT 2866
DB 2954 ACCCACTCCACACCTTGAGCTATATTGAGAGTGAAGTGAATGAATG---GTGGAGGGG 3010
QY 2867 AGCAGGAGAGGTGTTCTGTGGGTGAGAGGTGAGGGGAAAGCCAGAGCAGAGGAGTCTGG 2926
DB 3011 GGCTGGAGGGGTGTTCTGTGGGTGAGAGGTG-AGGGAGAGGCCAGAGCAGAGGAGTCTGG 3069
QY 2927 CTTTGTCTCTGAAACAATGTCTACTTATTATTAACAGGCACTGCTGTAAAGACCA 2986
DB 3070 CTTTGTCTCTGAAACAATGTCTACTTATTATTAACAGGCACTGCTGTAAAGACCTG 3129
QY 2987 ACATCTAGACCTCTGAAAGACAGAGCCCTGAGAGCAGGGGTGTTCTCTGAGCTTG 3046
DB 3130 ACACCTTACAGCCTCTG-AGAGATGAGCAGCTTGAAGATGGGAGCGGGGCTTGAAGGGC 3188
QY 3047 GGTGCTTGAATGAGGCAAAAGAGGSCATGAGTGAATGAAGGCTCCAGAGG- 3102
DB 3189 TAGGGTCTTGAAGACACCAAAAGAGGAATGAGATTAAGCCCTGAGAGGCTTGG 3248
QY 3103 -----GTTAGAAAGGCACTTGGAAAGGGTCACTGTCAGAGCCCTATCCATGAA 3156
DB 3249 AAAGGGCAACAATAATTCCTGCTGGAGAGGGGCAAGTGAAGAGCCCAACCCGATGAA 3308
QY 3157 TCTGAGCCTGGAG-CCAATGTGTATATCTCTGGGCTGTCAGAGGATTCMAAGCAGCA 3215
DB 3309 TGCAGTGCCTGGAGACCAAGATGTGTAAATCCCAAGGCTGCGCAG-----GCA 3356
QY 3216 CTTGCAATCCCTGTGAGAGCCTGGAGAGGCGGAAGGAGCAACCCCACTTATACCTTTTC 3275
DB 3357 CTTGCAACCTCTGTGAGAGCCTGGTGAAGGCAAGGAGCAGGCTCTACACCCCTCTCT 3416

```
QY 3276 TCCCTAGCCCCAGATTAAACCTCTGCGCTTCCCTTCCCATCCCATCAGAGTGG 3335
Db 3417 CTCTCAGGCCCTTAGGATAGCACCATTCTT-----CTGCCACCCCACCCGAGGACA 3471
QY 3336 GAGGGTTGCAGAGGAGGTAATAAACCTACATGTCCAAACATCATGTGCACGATATATG 3395
Db 3472 GAGGGTAGGACG--GGGAGGGCAAAAGCTTAGTGTGCATAATGCCGATGTGCA--AATGTGTG 3529
QY 3396 GATCAGTATGTGTAGAGCA--AGAAAGAAATCTGCAGGCTTAACTGGGTTAATGTGT 3452
Db 3530 CATCCATGTACACAAAGAAAGTAGAAAGAAATCTGCAAGCTTGAGTGGTAAAGTGTGA 3589
QY 3453 AAGTCGTGTGATGTGTGTGTGTGTGACTGAAAACGGGCATGGCTGTGCAGCTGTCA 3512
Db 3590 AAAATCTGCATGTGTGCTGAAGATGGCACAGACAGGTCAAG----- 3633
QY 3513 GTTCTGTGCTGAGTTACCACTGCAGAGTTGTGTGTAAATTGCCAAGCAAGCAAGTGG 3572
Db 3634 --TCTGATGTGAGAGTGCCTGAACCTGGGGTTCTGTGAAAATCTGCTGAGGCGGCAGG 3691
QY 3573 GTGAATCCCTTCATGTGTTAAAGATTTGATGATGGCTGCATCTCAAGACCATGGA 3632
Db 3692 GAGAAATCAGCTGCATTGCGTGCAGAGTGTGATGTGGCCACTATTC--AGGAGCATTTAG 3750
QY 3633 AAATGAAATGGACACTTATATGTGTCTTAAGCTAAGTAG--CAAGGTCTTTGGAGG 3689
Db 3751 GGAAGGGGTGGGACTTCACAGCTGTCCAAACAGGGTGGCTCAAGACCTTGGGAGA 3810
QY 3690 ACACTGTCTAGAGATGTGGGCAACAGAGACTACAGACATCTGTACAGATTAAGAG 3749
Db 3811 ACACTGTCTGAAAGACTTGGGGAACG-----AAGGAGACAGGCATGGCACTTATGCA 3865
QY 3750 AGAGAGAGGGGGGTGTAGAAATCTTACTATCAAAAGGAACTGATGTGTGCACTGCA 3809
Db 3866 ACTGAGGCCAGGACAGAA-----TTTCTGACAAAAGAAACTGAGCCATGAGATGGA 3919
QY 3810 AAGTGAATGCTCCCTAGACATCATGACTTTGTCTGTGGGAGCCAGCACTGTGGAAGT 3869
Db 3920 CAACAGATTCCTTCCCTGGGCA--CATACTGCAGCTTTTAGTCCCTAGCACTTGGGGCTC 3977
QY 3870 TCAGTCTGAGAGTAGAGAGGCTCCCTCAGCCTGAAGCTATGCAGATAGCCAGGGTTG 3929
Db 3978 CAGTACTAACAGAGGAAGATGCTCCAGCCTTGGGACTGTGTGAGGAAATGTCAGAAATGG 4037
QY 3930 AAAGG--GGGAAGGAGAGCTGCGATGGAGCTTGTGTGTTGGAGCGAGGGAAGATAT 3988
Db 4038 AAAGGAGAGGCTGGGGAACAGGGAGGAAAGCCCATGTGTTGGAGGCGGAGACAGGCAT 4097
QY 3989 TAAGCCTGGAAGAAGAGTGAACCTTACCCAGTTGTCAACCTCAACCTTCAGATTAAAA 4048
Db 4098 TTGGCCTGCGAGAGAGGTACCTCACCATGTGTTTC--AGTCTCACTTGGGGGAAAAA 4156
QY 4049 TAACTGAGTAAAGGC-----CTGGTAGGGGAGTGTGTGAGACGCTCTGTCTTCC 4103
Db 4157 TAACTGAGTAAAGGCCATGAGGAGGTGGGAGGCGGTGTGAGAAAGTCTGTCTTCC 4216
QY 4104 TC--TGCATGCCCTGAGGCCCTTTGGGAGGAGG--AATGTGCCAAGACTAAAAAAGG 4160
Db 4217 ACATATCTGTCTCATTCAGCCCTTTGAAGGGGAGGTAAATGTGCTTAAGGACTAAAAAAGG 4276
QY 4161 CCATGAGCCAGAGGGGCGAGGCAACAGACTTTTCATGGGCAAACTTGGGGGCC 4216
Db 4277 CCGTGAGCCAGAGAGGCTGGGGCAGACCTTTCTTGGGCAAAATCAGGGGGGCC 4332
```

This Page Blank (uspto)